Molecular and Evolutionary Studies of Race-specific Avirulence Factors from Wheat Powdery Mildew (*Blumeria graminis* f. sp. *tritici*)

Dissertation

zur

Erlangung der naturwissenschaftlichen Doktorwürde

(Dr. sc. nat.)

vorgelegt der

Mathematisch-naturwissenschaftlichen Fakultät

der

Universität Zürich

von

Kaitlin Mc Nally

aus

den Vereinigten Staaten von Amerika

Promotionskommission

Prof. Dr. Beat Keller (Vorsitz und Leitung der Dissertation)

Dr. Salim Bourras

Prof. Dr. Bruce McDonald

Prof. Dr. Christoph Ringli

Zürich, 2017

For

Mama and Daddy

"When I was five years old, my mother always told me that happiness was the key to life. When I went to school, they asked me what I wanted to be when I grew up.
I wrote down 'happy'. They told me I didn't understand the assignment, and I told them they didn't understand life." - John Lennon

Contents

Summary

Plant innate immunity is a complex system of preformed and inducible barriers. Inducible defense responses such as localized cell death are initiated by immune receptors which perceive the presence of pathogens by specific recognition of their secreted effector proteins. The characteristics of these so-called avirulence (*Avr*) factors from the pathogen and the basis of the specificity of their recognition by host immune receptors is not well understood. The topic of this thesis is the identification and characterization of *Avrs* from the obligate biotrophic fungal pathogen *Blumeria graminis* f. sp. *tritici* (*B.g. tritici*), which causes powdery mildew disease on wheat (*Triticum aestivum*).

B.g. tritici exhibits extreme specificity for its host, wheat, and is unable to colonize the nonhost plant *Arabidopsis thaliana*. However, *B.g. tritici* exhibits race-specific microcolony formation on an *Arabidopsis* triple mutant deficient in nonhost defense responses. A preliminary genome-wide association study of potential *Avrs* in *B.g. tritici* which might mediate this race-specific microcolony formation indicated multiple factors are contributing to the observed phenotype.

The main focus of this thesis was the identification and characterization of an *Avr* recognized by the *Pm3* allelic series from wheat conferring race-specific resistance against *B.g. tritici*. Phenotypes segregating in the progeny of a fungal genetic cross revealed three major loci in the pathogen, the first genetically interacting with all tested *Pm3* alleles and the second interacting specifically with the *Pm3a* and *Pm3f* alleles. We located candidate genes for these two loci by fine mapping. Functional validation in transient assays confirmed the cloning of $AvrPm3^{a2/f2}$, specifically recognized by *Pm3a* and *Pm3f*, and $SvrPm3^{a1/f1}$, the suppressor of recognition, demonstrating the *Avr-R-Svr* model of interaction in powdery mildews.

To further study the basis of recognition specificity of the *Pma/f* alleles for $AvrPm3^{a2/f2}$, we used an approach that combined two sources of sequence diversity information. A worldwide collection of mildew was examined for $AvrPm3^{a2/f2}$ allelic diversity and this information was combined with sequence diversity of the $AvrPm3^{a2/f2}$ effector family to design single-residue altered $AVRPM3^{A2/F2}$ constructs. Transient coexpression of these constructs with PM3A, PM3F, and $PM3F^{L456P/Y458H}$ (modified for enhanced signaling) led to the observation of a high frequency of disruptive mutations, and a domain where residue alterations enhanced the cell death response. This putative interaction domain shares no overlap with naturally polymorphic residues, suggesting $AvrPm3^{a2/f2}$ sequence diversity is not the result of selection pressure from recognition by *Pm3a/f*.

Finally, follow-up analyses comparing $AvrPm3^{a2/f2}$ and $SvrPm3^{a1/f1}$ gene expression, signatures of selection pressure, and diversity and functional properties of $AvrPm3^{a2/f2}$ natural polymorphisms between mildew populations revealed divergent strategies to overcome *Pm3a/f* resistance in each population. We hypothesize that evasion of recognition is not the main driver of $AvrPm3^{a2/f2}$ allelic diversification, but allelic diversity might rather be derived from novel effector function. This is supported by the observation of balanced selection in the genetically diverse host and pathogen populations of Israel, while gain-of-virulence mutations have likely resulted from intensive agricultural practices using host monocultures in the US. Therefore, we demonstrate that mechanisms to evade recognition differ between populations and are therefore likely dependent on population structure of both host and pathogen.

Zusammenfassung

Angeborene Immunität in Pflanzen ist ein komplexes System, bestehend aus verschiedenen Barrieren, einige sind konstitutiv, während andere induzierbar sind. Diese induzierbaren Abwehrreaktionen wie beispielsweise lokaler Zelltod werden durch Immunrezeptoren ausgelöst. Diese Rezeptoren erkennen die Anwesenheit eines Krankheitserregers (Pathogens) anhand von dessen spezifisch ausgeschiedenen Effektorproteinen. Die Besonderheiten dieser so genannten pathogenen Avirulenzfaktoren (*Avrs*), und die Grundlage ihrer spezifischen Erkennung durch Immunrezeptoren des Wirts sind nicht gut verstanden. Das Thema dieser Doktorarbeit ist die Identifizierung und Charakterisierung von Avirulenzfaktoren des obligat biotrophen Pilzes *Blumeria graminis* f. sp. *tritici* (*B.g. tritici*), dem Auslöser der Mehltaukrankheit in Weizen (*Triticum aestivum*).

B.g. tritici weist eine starke Wirtsspezifität für Weizen auf und kann deshalb *Arabidopsis thaliana* normalerweise nicht befallen. *B.g. tritici* kann jedoch, abhängig von der Mehltaurasse, Mikrokolonien auf *Arabidopsis* Pflanzen bilden, welche in drei Immunitätsfaktoren mutiert sind. Eine GWAS (genome-wide association study) Studie potentieller Avirulenzfaktoren in *B.g. tritici*, welche diese rassen-spezifische Mikrokolonisierung vermitteln, ergab, dass mehrere Faktoren für diesen Phänotyp verantwortlich sind.

Der Hauptfokus dieser Arbeit liegt auf der Identifizierung und Charakterisierung eines Avirulenzfaktors welcher durch die *Pm3* Allele in Weizen erkannt wird. Diese *Pm3* Allele kodieren für Immunrezeptoren und vermitteln rassen-spezifische Resistenz gegen *B.g. tritici*. Die Analyse segregierender Phänotypen in den Nachkommen einer Kreuzung unterschiedlicher Pilz-rassen führte zur Identifikation von drei genetischen Regionen (Loci) im Pathogen. Während der erste Lokus mit allen getesteten *Pm3* Allelen interagierte, war der zweite Lokus spezifisch für die Allele *Pm3a* und *Pm3f*. Mit Hilfe von "fine-mapping" Methoden wurden mehrere Kandidaten-gene innerhalb dieser Loci identifiziert. Eine funktionelle Validierung dieser Gene mit Hilfe transienter Expressionsmethoden bestätigte, dass *AvrPm3*$^{a2/f2}$ spezifisch durch *Pm3a* und *Pm3f* erkannt wird, wohingegen *SvrPm3*$^{a1/f1}$ diese Erkennung unterdrückt. Diese Interaktion führte zur Entwicklung des *Avr-R-Svr* Interaktionsmodels in verschiedenen Gräsermehltauarten.

Um weitere Einblicke in die Grundlagen der Erkennungsspezifität der *Pm3a/f* Allele für *AvrPm3*^{a2f2} zu erhalten, wurde die Diversität der *AvrPm3*^{a2f2} DNA Sequenz mit zwei Ansätzen untersucht. Zum einen wurde eine Kollektion von verschiedenen weltweiten Mehltauisolaten hinsichtlich ihrer *AvrPm3*^{a2f2} Allelvielfalt analysiert. Diese Information wurde mit einer Analyse der Sequenzdiversität der *AvrPm3*^{a2f2} Effektorproteinfamilie kombiniert, um Expressionskonstrukte zu konstruieren, welche sich in einer Aminosäure unterscheiden (AVRPM3$^{A2/F2}$). Die transiente Ko-expression dieser Konstrukte zusammen mit PM3A, PM3F und PM3F$^{L456P/Y458H}$ (modifiziert um verstärkte Signaltransduktion zu vermitteln) zeigte, dass viele dieser Mutationen zu einem Verlust der Erkennung zwischen Immunrezeptor und Effektorprotein des Pathogens führen. Des Weiteren wurde eine Sequenzdomäne entdeckt, welche, wenn mutiert, zu einer verstärkten Zelltodantwort führt. Diese Domäne überlappt nicht

mit der Region im Avr Protein welche natürlich vorkommende Polymorphismen enthält. Daher ist es wahrscheinlich, dass die *AvrPm3*^{a2f2} Sequenzdiversität nicht durch den Selektionsdruck entsteht, der sich aus der Erkennung durch *Pm3a/f* ergibt.

Ein Vergleich zwischen der Genexpression von *AvrPm3*$^{a2/f2}$ und *SvrPm3*$^{a1/f1}$, der Selektions-drucksignaturen, sowie der Vielfalt und funktionellen Eigenschaften von natürlichen *AvrPm3*$^{a2/f2}$ Polymorphismen in verschiedenen Mehltaupopulationen weltweit offenbarte unterschiedliche Strategien im Pathogen, um die *Pm3a/f* Resistenz zu überwinden. Wir vermuten, dass die Überwindung der Erkennung durch das Resistenzgen nicht der Hauptgrund für die *AvrPm3*$^{a2/f2}$ Allelvielfalt ist. Vielmehr ist wahrscheinlich, dass sie durch eine neue Funktion des Effektor-proteins entstanden ist. Dies zeigt sich am Beispiel der ausgeglichenen Selektion zwischen Wirt- und Pathogenpopulationen aus Israel. Wirtsmonokulturen aus den USA hingegen, welche durch intensive Landwirtschaft entstanden sind, zeigen eine starke Anreicherung an virulenten Mutationen. Die Mechanismen, welche zur Überwindung der Erkennung führen, unterscheiden sich in den verschiedenen Populationen. Daher vermuten wir, dass sie abhängig von der Populationsstruktur des Wirts und des Pathogens sind.

CHAPTER 1

General Introduction

1.1 Immunity in plants

In the course of evolution plants have developed a multi-faceted immune response that prevents pathogen invasion and disease. This innate immunity can be divided into three mechanisms: (1) extracellular receptor-like kinases (RLKs) and receptor-like proteins (RLPs) that perceive pathogen associated molecular patterns (PAMPs) or pathogen effectors in the apoplast; (2) intercellular nucleotide-binding leucine rich-repeat receptors (NLRs) that sense the presence or activity of secreted effectors and result in race-specific, major gene resistance; and (3) quantitative resistance, which is characterized by a partial resistance conferred by the function of multiple genes and that is generally more effective against a broader spectrum of pathogens (Jones and Dangl, 2006; Zipfel et al, 2014; Krattinger and Keller, 2016). The contribution of major resistance genes to immunity against adapted pathogens is well-studied, as it is an important source of resistance in plant breeding. The partial contribution of multiple quantitative genetic factors in immunity is harder to characterize but the complexity of this kind of immune response constitutes an effective barrier to infection by non-adapted pathogens.

1.2 Non-host resistance

Plant immunity is therefore divided into two categories with overlapping molecular mechanisms: 'host resistance' against adapted pathogens and 'non-host resistance' against non-adapted pathogens. Non-host resistance consists of both pre- and post-invasion defense mechanisms (Lipka et al, 2005). Pre-invasion defenses include the physical barriers constitutively present in plants, such as the waxy cuticle or extra-cellular matrix. It also includes inducible physical barriers, which are formed in response to the presence of pathogens, such as toxic compounds which are secreted and callose deposits which form structures called papillae beneath attempted cell-penetration sites. Forward genetics studies in *Arabidopsis* have been very successful for identifying the genetic factors underlying pre-invasion defense responses (Collins et al, 2003). Mutations in genes involved in pre-invasion responses lead to a partial breakdown of non-host defense responses and are therefore useful for identifying additional factors involved in the resistance against non-adapted pathogens.

1.3 Resistance against adapted-pathogens

While pre-invasion defenses are not considered obstacles to adapted pathogens, the components of post-invasion responses are crucial for both host and non-host resistance since they provide immunity to any pathogens which are able to bypass pre-invasion defenses. Post-invasion responses involve the action of the RLKs, RLPs and NLRs that either recognize PAMPs or pathogen secreted effectors, inducing PAMP-triggered immunity (PTI) or effector-triggered immunity (ETI), respectively (Jones and Dangl, 2006). Post-invasion responses depend upon effective signaling of these receptors. Combining mutations that disrupt these signaling pathways with mutations in pre-invasion components has been shown to severely compromise non-host immunity (Lipka et al, 2005). This highlights the importance of ETI not only for preventing disease by adapted-pathogens but also for immunity against non-adapted pathogens.

1.4 Filamentous fungal and oomycete pathogen effectors

Pathogens secrete effectors to silence host defense responses and manipulate the cellular metabolism. Effectors are typically small, secreted proteins with no homology to known proteins. Filamentous fungal and oomycete pathogen effectors are typically members of structurally conserved but sequence-unrelated families (Nemri et al., 2014; de Guillen et al., 2015; Bourras et al., 2015; Spanu 2017), in fast evolving genomic regions (Grandaubert et al., 2014; Huang et al., 2014). Fungal effectors have been shown to have many different targets in plants such as RLKs/RLPs, proteases, salicylic acid signaling components, microtubules, vesicles and metabolic pathways (Giraldo and Valent, 2013; Lo Presti et al., 2015). Once effectors are recognized by host-derived resistance (*R*) genes thereby eliciting defense responses, they are referred to as avirulence (*Avr*) factors.

1.5 The study of recognition specificity

The specificity of *R* genes for their cognate *Avrs* results in what is called race-specific resistance. A resistance gene is only effective against the races of a pathogen which employ the cognate AVR protein during infection. Determining the basis of recognition specificity of a receptor is an important topic for the development of improved resistance for agriculture. Traditionally it was thought that R proteins recognize their cognate pathogen-encoded Avrs in a gene-for-gene manner (Flor, 1971). Gene-for-gene interactions have been extensively characterized in the resistance of flax (*Linum usitatissimum*) to the flax rust pathogen (*Melampsora lini*), where single *R* genes and their alleles (*L, M*, and *P*) recognize specific *Avr* genes (*AvrL567, AvrM, AvrP, AvrP123*, and *AvrP4*) from the fungus (Dodds et al, 2006, Ellis et al, 2007; Catanzariti et al, 2010).

However, recently it has been demonstrated that some interactions can be much more complex and involve several factors, such as a suppressor of recognition (Bourras et al, 2015 [Chapter 3]). Receptor specificity can therefore be altered by the action of additional interacting factors, such as effector modifiers (Boehnert et al, 2004) or suppressors of recognition (Bourras et al, 2015 [Chapter 3]).

There are an increasing number of mechanisms by which NLRs function to recognize effectors and induce signaling that have been described. Traditionally it was thought that NLRs directly bind to effectors through the highly polymorphic leucine-rich repeat domain for which there are many examples (Jia et al, 2000; Ellis et al, 2007; Krasileva et al, 2010). More recently integrated domains mimicking effector targets have been identified as the primary site of interaction in several NLRs (Maqbool et al, 2015; Ortiz et al, 2017). NLRs have also been shown to indirectly recognize effectors by detecting modifications to their effector target, called a 'guardee', or a host protein mimicking the target, called a 'decoy' (van der Hoorn and Kamoun, 2008).

1.6 Utilizing natural diversity as a tool to study specificity

A useful strategy for the study of recognition specificity has been to characterize the natural diversity of *R* and *Avr* alleles. *Avrs* are thought to be under strong selection pressure to evade recognition by their cognate *R* genes. The natural diversity of *Avr* alleles can provide important clues to the nature of *Avr-R* interaction. For example, if an effector is dispensable or recognition occurs indirectly through its function, then inactivation or deletion of the gene is often observed (Dangl and McDowell, 2006). This is the case for *Avr-Pii* and *Avr-Pia* from *Magnaporthe oryze*, the causal agent of rice blast, where presence/absence polymorphisms determine virulence on *Pii* and *Pia* containing rice cultivars (Huang et al, 2014). In other cases an effector important for pathogen fitness needs to be maintained by the pathogen. Selection to maintain effector function leads to mutations in the protein sequence that are likely to disrupt recognition (Chen et al, 2014; Blondeau et al, 2015). This was observed for *AvrLm2* from *Leptosphaeria maculans,* where *AvrLm2* was amplified from all isolates screened, indicating its importance for pathogen fitness, and a single polymorphism in the gene associates with avirulence phenotypes on *Lm2* containing canola (*Brassicus napus*) cultivars (Ghanbarnia et al, 2015). In both cases, the characteristics of the natural diversity reflect the dispensability of the effector and its evolutionary potential to evade recognition.

Protein structure of avirulence genes can also be informative for interpreting the effects of polymorphisms in the natural diversity on recognition. Crystal structures of two members of the *AvrL567* family (*A* and *D*) from flax rust (*Melampsora lini*) revealed that most polymorphisms in the effector family map to surface residues and that a subset of these are likely to determine specificity of recognition by the *L5* and *L6* resistance

genes, which was confirmed *in vitro* (Wang et al, 2007). Predicting the effects of polymorphisms using information from structural models can be used to design studies of recognition specificity as well as for interpreting the results.

1.7 Natural diversity in the context of coevolution

It is crucial for the correct interpretation of the role of natural diversity in recognition specificity to take into consideration the co-evolutionary dynamic between *Avr* and *R* genes. AVR-R interactions often exhibit so-called 'arms race' or 'trench warfare' dynamics characterized by reciprocal evolution or frequency-dependent selection, respectively. Many examples to date support the theory that arms race dynamics are derived from direct AVR recognition by the host R protein, causing diversifying selection and higher genetic diversity in both *R* and *Avr* genes (Leonelli et al, 2011; Kanzaki et al, 2012; Dodds et al, 2006). Recognition specificity of the direct interaction between *Avr-Pik* alleles from *M. oryzae* and the *Pik* alleles from rice (*Oryzae sativa*) associate precisely with natural sequence diversity of both proteins and to their subsequent binding affinities (Kanzaki et al, 2012). In this instance, associating the natural diversity with specificity revealed the reciprocal co-evolution between ligand and receptor. Conversely, trench warfare is often associated with indirect recognition of effector function, which is characterized by balancing selection and low genetic diversity (Van der Hoorn et al, 2002; Jones and Dangl, 2006; Xiao et al, 2008). A classic example of balancing selection is the *Rpm1* resistance gene in *Arabodopsis*, which is divided into two allelic classes of resistant and susceptible alleles that diverged in the ancient past and are maintained at comparable frequencies (Stahl et al, 1999). These co-evolutionary dynamics provide clues to the nature of the *Avr-R* interaction.

1.8 The powdery mildew disease

The powdery mildew disease of wheat (*Triticum aestivum*) is caused by the obligate biotrophic fungal pathogen, *Blumeria graminis* f. sp. *tritici* (*B.g. tritici*). Powdery mildews exhibit extreme host specificities, where distinct *formae speciales* infect specific cereal hosts, such as wheat, barley (*Hordeum vulgare*), oats (*Avena sativa*), and rye (*Secale cereale*) (Troch et al, 2014). *B.g. hordei* (barley powdery mildew) and *B.g. tritici* are estimated to have diverged 6.3 (±1.1) million years ago, following which *B.g. tritici* experienced a dramatic increase in its effector repertoire, possibly to accommodate its new host specificity (Wicker et al, 2013). *B.g. tritici* can infect both wild and domestic *Triticum* species (Ben-David et al, 2016). The modern, agriculturally relevant hosts, *T. aestivum* and *T. durum* (bread wheat and durum wheat, respectively) were domesticated in the Fertile Crescent approximately 10,000 years ago (Özkan et al, 2002; Peng et al, 2011). The wheat powdery mildew disease spread along with wheat cultivation to Europe, Asia, and more recently across the oceans to the United States

and Australia. It is capable of long distance dispersal documented at up to 650 km (Hermansen et al, 1978; Wolfe and Schwarzbach, 1978), but is strongly limited by geographical and climatic factors, resulting in instances of well mixed as well as strongly divided population structures (Ben-David et al, 2016; Cowger et al, 2016). No one has yet investigated how these different population dynamics affect the evolution of *Avr-R* interactions between powdery mildew and wheat.

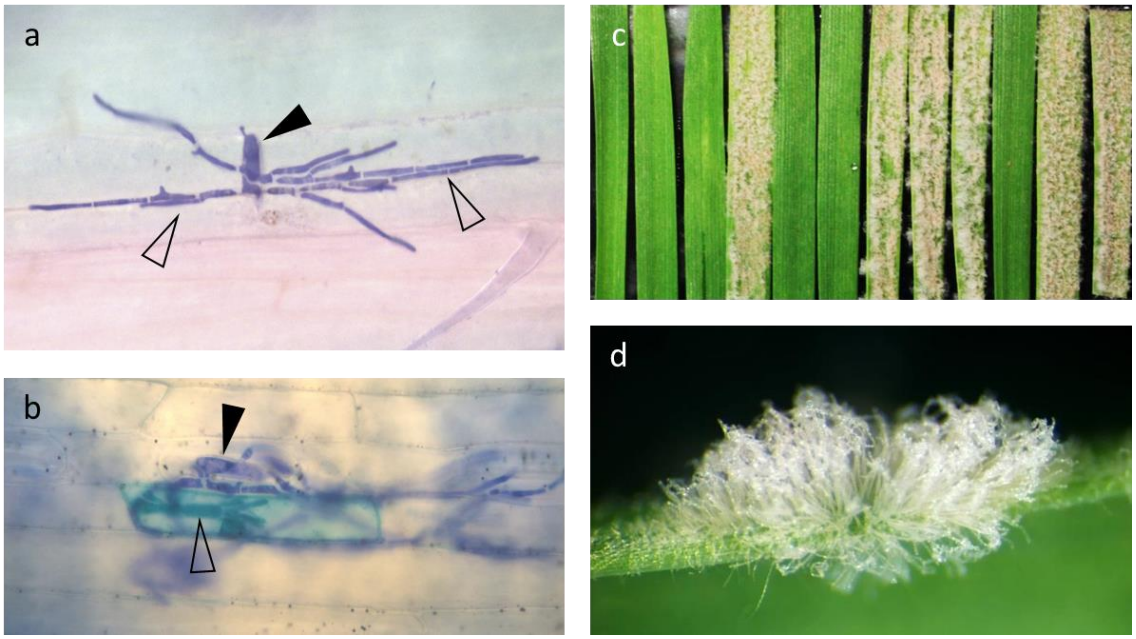


Figure 1. *Blumeria graminis* f. sp. *tritici*, causal agent of wheat powdery mildew. (a) Coomassie blue stained spore (black arrowhead) with secondary hyphae (outlined arrowheads, photo courtesy of Stukenbrock et al, 2016). (b) Coomassie blue stained spores (black arrowhead) and GUS stained wheat cell with developed haustoria inside (outlined arrowhead). (c) Phenotyping of a mildew isolate on differential wheat lines carrying resistance genes against powdery mildew. (d) A powdery mildew pustule on a wheat leaf.

1.9 The *Pm3* race-specific allelic series

Pm3 from wheat forms an allelic series of genes encoding NLRs and is an important source of race-specific resistance against *B.g. tritici* (Brunner et al, 2010). The seventeen true alleles of *Pm3* (*a-g, k-t*) cloned from hexaploid and tetraploid (*Triticum durum*) wheat cultivars and land races share particularly high sequence identity (>97%) (Yahiaoui et al, 2004; Srichumpa et al, 2005; Bhullar et al, 2009; 2010). Several of these alleles diverged in the recent past (Yahiaoui et al, 2009), the most recent of which (*a, b, c* and *f*) likely evolved through gene conversion events leading to large polymorphic blocks in the NB and LRR domains (Sela et al, 2014). The *Pm3a* and *Pm3f* alleles share polymorphic sequence blocks distinct from other alleles, but differ mostly in their nucleotide-binding ARC1 and ARC2 domains (Brunner et al, 2010). As the *Pm3a* and

Pm3f alleles share overlapping recognition spectra toward powdery mildew races, where the races recognized by *Pm3a* include those recognized by *Pm3f* (Brunner et al, 2010), it was hypothesized that these mutations were responsible for PM3A being a 'stronger' allele than PM3F. Later it was demonstrated that two mutations in this region of PM3F towards residues in PM3A were sufficient to enhance the autoactive cell death response in *Nicotiana benthamiana* transient protein expression assays to levels comparable to PM3A (Stirnweis et al, 2014a).

In comparison, the *Mla* allelic series in barley against barley powdery mildew (*Blumeria graminis* f. sp. *hordei*) and the *L* alleles from flax also constitute true allelic series, but the alleles are more sequence divergent than *Pm3* (Seeholzer et al, 2010; Ravensdale et al, 2012). The *L5* and *L6* alleles share overlapping specificities for *AvrL567* alleles, but the larger spectrum of *L6* does not entirely include the spectrum of *L5* (Ravensdale et al, 2012). Conversely, it was found that AVR_{a1} and AVR_{a13} of the *Mla1* and *Mla13* alleles are sequence-unrelated *Avr* genes (Lu et al, 2016). Given the recent evolution and overlapping specificities of the *Pm3* alleles, it is difficult to predict what kind of interaction will be observed for the *AvrPm3-Pm3* system. The relatively recent evolution of the *Pm3* alleles suggests that *AvrPm3-Pm3* interactions may have been established in the relatively recent past. Together this provides a unique opportunity to study the basis of recognition specificity of a recently diverged multi-allelic resistance gene and potentially gain insights into the coevolution of the highly specific interaction of wheat and the powdery mildew pathogen.

1.10 Functional studies of *Blumeria* effectors

In addition to recognition specificity, an important topic of study in *Blumeria* ff. spp. is effector function. Previously, 491 effectors were identified in *B.g. hordei*, these grouped into 72 families and where shown to be predominantly expressed in haustoria (Pedersen et al, 2012). Studies into the specific functions of several effectors including CSEP0055, CSEP0105, and CSEP0162 revealed the individual importance of these genes for fungal aggressiveness, and demonstrated interactions with specific host targets (Zhang et al, 2012; Ahmed et al, 2015). Additional studies using host-induced gene silencing revealed the importance of eight effectors for virulence of *B.g. hordei*, and two of these proteins, BEC1011 and BEC1054, were verified by complementation assays to be bona fide effectors interfering with host cell death processes (Pliego et al, 2013); however, the exact function of these effectors remains undetermined. Interestingly, many effectors including two of the eight identified in the previous study as well as *Avrs* and *Svrs* from *B.g. tritici*, share structural homology to ribonucleases (Pedersen et al, 2012; Pliego et al, 2013; Bourras et al, 2015 [Chapter 3]; Lu et al, 2016; Praz et al, 2017). It has been hypothesized that these genes all share a common ancestor that existed before the divergence of *hordei* and *tritici* ff. spp., and it has been suggested that

they may play a role in RNA trafficking during infection (Pedersen et al, 2012; Spanu et al, 2017). Many other effectors share no homology to any known proteins, and these will also be interesting topics of further study to elucidate the varied functions of *B. graminis* effectors during infection.

1.11 Aims of the thesis

The availability of the *Arabidopsis pec* triple mutant deficient in non-host immune responses, allowing differential microcolony formation by races of *B.g. tritici*, makes possible the study of the basis of race-specific resistance in a non-host interaction. Here we initiate two approaches to study the basis of this specificity on the pathogen side, both by genome-wide association studies and a genetic cross of mildew isolates to map quantitative trait loci in the pathogen. Preliminary results suggests that multiple factors on the pathogen side are responsible for race-specific virulence on the *pec* triple mutant, and that precisely quantified isolate phenotypes will be necessary for further characterization of this interaction.

The main goal of this work was to identify and characterize the factors involved in recognition by the *Pm3a* and *Pm3f* alleles. To this end the progeny of two genetic crosses of isolates with polymorphic phenotypes on *Pm3* alleles *a-f* were analyzed for segregating phenotypes and multiple genetically interacting loci were identified in the fungus. We cloned *AvrPm3*$^{a2/f2}$ and functionally verified recognition by *Pm3a* and *Pm3f* in transient assays. Additionally, we identified a suppressor of recognition, *SvrPm3*$^{a1/f1}$, and functionally verified its quantitative role in suppressing recognition by *Pm3f*. Based on our findings, we proposed the *Avr-R-Svr* model of interaction for the recognition of powdery mildew by *Pm3* alleles.

To further study the basis of recognition specificity of *Pm3a/f* for *AvrPm3*$^{a2/f2}$, we describe the natural diversity in a worldwide collection of mildew isolates. Based on this diversity and using also the diversity from the sequence divergent and structurally related *AvrPm3*$^{a2/f2}$ effector family, we designed 85 single residue alterations across the AVR protein and analyzed their effects on the cell death response by PM3A and PM3F in transient assays. The information from the transient assays was compared to studies of selection pressure in the effector family and putative domains were identified as having a role in recognition whereas other domains were implicated in effector function. Additionally, preliminary analyses of potential structural models were conducted, based on the conserved structure of the sequence divergent *AvrPm3*$^{a2/f2}$ effector family. A comparison of the patterns of natural diversity of the AVRPM3$^{A2/f2}$ variants among distinct populations revealed population specific coevolution of the *AvrPm3*$^{a2/f2}$*-Pm3a/f* interaction.

CHAPTER 2

Two Approaches to Identify Pathogen-Encoded Factors Determining Quantitative Non-Host Defense Responses in *Arabidopsis* against *Wheat powdery mildew*

Kaitlin Elyse McNally[1], Coraline Praz[1], Fabrizio Menardo[1], Linda Lüthi[1,2], Helen Zbinden[1], Tina Jordan[3], Salim Bourras[1] and Beat Keller[1]

[1]Institute of Plant and Microbial Biology, University of Zürich, Zollikerstrasse 107, 8008 Zürich, Switzerland

[2]Present address: Institute of Pathology and Molecular Pathology, UniversityHospital Zürich, Schmelzbergstrasse 12, 8091 Zürich, Switzerland;

[3]Center for Plant Microbiology, Universität Tübingen, Auf der Morgenstelle 32, 72076 Tübingen, Germany.

Unpublished data

Abstract

In nature, nonhost resistance prevents wheat powdery mildew (*Blumeria graminis* f. sp. *tritici*) from infecting the non-adapted host, *Arabidopsis*. Recently it was demonstrated that three mutations in resistance related genes in *Arabidopsis* (*pen1, eds1, campta3-4D*) are sufficient to partially inactivate this nonhost resistance, allowing *B.g. tritici* to penetrate the leaf surface and form non-sporulating microcolonies in a race-specific manner. In this study we seek to understand genetic factors determining race-specific virulence of *B.g. tritici* isolates on the *Arabidopsis pec* triple mutant. We visualize single spore interactions microscopically to characterize disease progression. In addition, we explored two approaches to genetically characterize virulence factors: genome-wide association studies (GWAS) and a classic genetic cross. None of the candidates identified by GWAS are secreted effectors, suggesting that physiological factors rather than effector recognition might underlie race-specific virulence. We examined the genes implicated in melanization and identify candidate polymorphisms to explain differences in melanization, and ultimately, virulence. We crossed a highly melanized isolate with the highest frequency of microcolony formation to a less melanized isolate that is not able to penetrate the *pec* mutant, laying the groundwork for future investigations toward understanding the basis of race-specific resistance in the nonhost interaction of wheat powdery mildew and *Arabidopsis.*

2.1 Introduction

When all members of a plant species are immune to all isolates of a microbe that is pathogenic on other plant species, this is termed 'non-host' resistance (Gill et al, 2015). Although the mechanisms of resistance are the same in host and non-host interactions, there appears to be a correlation between effector-triggered nucleotide binding leucine rich repeat receptor (NLR) -based immunity in host interactions, and pathogen-associated molecular pattern (PAMP) –triggered receptor-like kinase and receptor-like protein (RLK/RLP) -based immunity in nonhost interactions (Schulze-Lefert and Panstruga, 2011). However, there are many examples of the role of effector recognition in nonhost interactions. For example, infection of various nonhost species by the fungal rice pathogen *Magnaporthe oryzae* ff. spp. is prevented by the presence of one to three genes from the pathogen (Yaegashi et al 1978; Valent et al, 1991; Tosa et al, 2006). In bacteria, an effector indistinguishable from *AvrA* cloned from *Pseudomonas syringae* pathovar (pv.) *glycinea* was found in all tested isolates of *P. syringae* pv. *tomato*, and is thought to be responsible for nonhost immunity of soybean to these isolates (Staskawicz et al, 1987). Finally, a single dominant locus conferring nonhost resistance to several *Albugo candida* ff. spp. in Arabidopsis was found to contain *WRR4*, an intracellular NLR (Parker et al, 1996; Borhan et al, 2008). Transgenic expression of *WRR4* in the adapted hosts conferred resistance to their host-adapted *ff. spp*. (Borhan et

al, 2010). Although well-studied, RLK/RLP and NLR-based resistances do not explain all instances of nonhost resistance. Quantitative resistance has also been suggested to be effective in nonhost interactions. Two examples include the three quantitative trait loci (QTLs) conferring nonhost resistance of lettuce to the downy mildew *Bremia lactucae* (Zhang et al, 2009), as well as overlapping barley resistance QTL to adapted and non-adapted species (or *formae speciales*) of the rust fungus, *Puccinia hordei* (Jafary et al, 2008; Dido et al, 2016).

A system that has been used extensively to study the molecular basis of nonhost resistance is the powdery mildew – *Arabidopsis* pathosystem. Powdery mildews infect a wide range of hosts in an extremely host-specific manner (Braun, 2011). The species that infects cereal grasses, *Blumeria graminis*, is divided into *formae speciales* with distinct host specificities (Troch et al, 2014). *B.g. tritici* is host-adapted to wheat (*Triticum aestivum*) and is very closely related to *B.g. hordei* that is adapted to barley (*Hordeum vulgare*) (Oberhaensli et al, 2011). In compatible interactions, *B.g. tritici* and *B.g. hordei* forcibly penetrate the host cell wall and the immature papillae to establish haustoria in close contact with the host plasma membrane. They then secrete effectors into the host cell via the haustoria to silence defense responses and reprogram the cell to allow for nutrient uptake. Once spores are able to feed on the nutrients from the host, they extend secondary hyphae across the leaf surface and establish secondary haustoria, thus producing thousands of conidia per pustule and completing their reproductive cycle in 5 days. In contrast, when *Blumeria* conidia are inoculated onto the nonhost *Arabidopsis*, they germinate and produce appressoria, but are unable to penetrate the leaf surface (Lipka et al, 2005).

Forward genetics studies using the powdery mildew – *Arabidopsis* pathosystem successfully identified three genes (*PEN1-3*) that when individually mutated allow penetration by non-adapted *B.g. hordei* (Collins et al, 2003). *PEN1* is a syntaxin that accumulates at attack sites with the vesicle-associated membrane protein (VAMP)721/722 and N-ethylmaleimide-sensitive factor (NSF) adaptor protein (SNAP)33, forming a SNARE complex which mediates vesicle fusion (Collins et al, 2003; Assaad et al, 2004; Kwon et al, 2008; Kwaaitaal et al, 2010). This complex is involved in the secretion of exosomes carrying callose to the site of attempted penetration to form the papillae. The second penetration resistance pathway involves *PEN2*, a peroxisome-localized myrosinase that hydrolyzes indole glucosinolates (Lipka et al, 2005). *PEN2* works together with *PEN3*, an ABC transporter, to metabolize and transport tryptophan-derived secondary metabolites to the site of infection (Stein et al, 2006; Bednarek et al, 2009). Additionally, the *PEN* genes have been shown to have other functions besides those involved in cell-wall defense (Kobae et al, 2006; Stein et al, 2006; Zhang et al, 2007, 2008; Schlaeppi and Mauch, 2010; Eisenach et al, 2012). Together, the *PEN* genes constitute 'pre-penetration' defenses. Double mutant studies

with the *PEN* genes found that combining mutations in these two pathways was sufficient to allow up to 90% penetration success but no post-invasive growth by *B.g. hordei* (*pen1-pen2* double mutants), and a delayed hypersensitive response (HR) (Lipka et al, 2005). Another recent study of *PEN* double mutants suggests that this delayed cell death response is the result of compromised effector recognition (Johansson et al, 2014).

Successful effector recognition depends on intercellular receptor signaling. Up to three pathways have been proposed to contribute to effective NLR signaling (Maekawa et al, 2011), but the best characterized involves the EDS1/SAG101 and EDS1/PAD4 complexes (Feys et al, 2001; Wiermer et al, 2005; Rietz et al, 2011; Wagner et al, 2013; Li et al, 2015). EDS1 is a lipase-like protein that is a major regulator of basal resistance and salicylic acid accumulation (Wiermer et al, 2005). It forms a complex with the potentially redundant SAG101 and PAD4 (Feys et al, 2005) and has been shown to be crucial for Toll-Interleukin Repeat (TIR-) TIR-NLR signaling, but not Coiled-Coil (CC-) NLR signaling (Yun et al, 2003; Maekawa et al, 2011). 'Post-penetration' defense involves signaling through the EDS1/SAG101/PAD4 complexes, which causes salicylic acid accumulation and eventually a hypersensitive response (HR) leading to cell death which halts pathogen invasion. The *pps* mutant contains mutations in the EDS1/SAG101/PAD4 signaling complexes combined with *pen* mutations (*pen2-pad4-sag101*) and was shown to allow *B.g. hordei* to form microcolonies by producing secondary hyphae up to 30% of the time, and even to allow the occasional production of conidiospores (Lipka et al, 2005). This represents a major breakdown of nonhost resistance by mutation of only three plant-encoded genes.

Recent studies using another immune deficient triple mutant, *pec* (*pen1, eds1, camta3-4D)*, show a similar breakdown of nonhost defenses in response to cereal powdery mildews (Jordan et al, in preparation). The *pec* mutant is deficient in key pre- and post-invasion defense related genes, and has an additional mutation in the CAMTA3 transcription factor (*camta3-4D*) which was reported to constitutively silence defense responses (Galon et al, 2008). While different *B.g. hordei* isolates show equal virulence capabilities on the *pec* and *pps* mutants, *B.g. tritici* isolates show an increased ability to form microcolonies on the *pec* mutant background (Jordan et al, in preparation). Not only do they preferentially grow on the *pec* mutant compared to *pps*, but remarkably, the immune-deficient *pec* mutant background revealed race-dependent differences in virulence phenotypes of *B.g. tritici* isolates in this nonhost interaction (Jordan et al, in preparation; this study). We hypothesize that race-specific genetic factors in *B.g. tritici* contribute toward overcoming the remaining nonhost defense responses in *Arabidopsis pec* in a quantitative manner.

In this study we described the observed responses of the *Arabidopsis pec* mutant to attack by non-adapted powdery mildews. Two approaches were initiated to further characterize the pathogen-encoded factors involved in this interaction. First, a GWAS using a selection of isolates with different genetic backgrounds was explored. Second, as genetic analysis has been a useful tool to study factors that determine compatible interactions in *Blumeria* (Parlange et al, 2015; Bourras et al, 2015 [Chapter 3]; Praz et al, 2017), we screened for the most virulent and avirulent isolates to choose parents of a biparental population. This cross should enable mapping of the factors governing the frequency of microcolony formation in this nonhost interaction. Finally, we explored polymorphisms in key melanin biosynthesis genes in the parental isolates as a possible explanation for their virulence phenotypes on the *pec* mutant.

2.2 Results

Infection by non-adapted powdery mildew is arrested at distinct stages in a race-specific manner

43 isolates from an international *B.g. tritici* and *B.g. triticale* collection (Europe, China, Israel, and the USA) with diverse geographic origins were selected for infection on the nonhost immune deficient *Arabidopsis pec* mutant (Table 1). Fresh spores were inoculated onto detached *Arapidopsis pec* leaves and observed 7 days post-inoculation (dpi). No macroscopic symptoms were visible, so cleared leaves were stained with Coomassie Blue to visualize fungal hyphal development microscopically. Microscopic observation revealed a variety of stages in which infection was stopped in a nonhost interaction of several *B.g. tritici* isolates with the immune deficient *Arabdopsis pec* mutant (Figure 1). Visual observations revealed differences in microcolony size between isolates, indicative of quantitative differences in virulence capabilities, however, microcolony size also occasionally varied between replicates of the same isolate, reflecting the environmental dependency of this phenotype. The not sufficiently controllable environment prevented reliable measurement of microcolony size. Microscopic visualization revealed autofluorescent defense responses in *Arabidopsis* to challenge by *B.g. tritici* (Figure 1A-F), including the deposition of callose and phenolic compounds at the penetration site. The callose deposition response under the first, second, and third contact sights characteristic of papillae formation was visualized by aniline blue staining (Figure 1A). Papillae were observed at all contact sites, appearing to be the most effective obstacle to successful infection (Figure 1B). Visual observations revealed that once papillae-based defense was overcome, haustorial encasement prevented further disease progression by *B.g. tritici* (Figure 1C), as was previously described by Meyer and colleagues (2009) for *B.g. hordei*. Spores which successfully penetrated the papilla occasionally formed a complete haustorium but nevertheless triggered a cell death response that is visible as whole-cell autofluorescence (Figure 1D,

Table 1. Isolates phenotyped on the Arabidopsis *pec* mutant.

Isolate	Origin	*Forma specialis*	PEC
CAP9-A1	CH	*B.g. triticale*	A
T3-6	CH	*B.g. triticale*	A
T3-9	CH	*B.g. triticale*	A
T4-7	CH	*B.g. triticale*	A
T5-13	CH	*B.g. triticale*	A
BU-18	CH	*B.g. triticale*	I
T1-23	CH	*B.g. triticale*	I
THUN-12	CH	*B.g. triticale*	I
7004	CH	*B.g. tritici*	A
07230	CH	*B.g. tritici*	A
94202	CH	*B.g. tritici*	I
07250	CH	*B.g. tritici*	I
96229	CH	*B.g. tritici*	I
96224	CH	*B.g. tritici*	V
10-8	CN	*B.g. tritici*	A
11-99	CN	*B.g. tritici*	A
12-50	CN	*B.g. tritici*	A
13-50	CN	*B.g. tritici*	A
14-17	CN	*B.g. tritici*	A
15-9	CN	*B.g. tritici*	A
41-5	CN	*B.g. tritici*	A
5-83	CN	*B.g. tritici*	A
1-25	CN	*B.g. tritici*	I
3-53	CN	*B.g. tritici*	I
6-6	CN	*B.g. tritici*	I
9-2	CN	*B.g. tritici*	I
36-70	CN	*B.g. tritici*	V
46-30	CN	*B.g. tritici*	V
6-69	CN	*B.g. tritici*	V
Isr15	IS	*B.g. tritici*	A
Isr215	IS	*B.g. tritici*	A
Isr217	IS	*B.g. tritici*	A
Isr208	IS	*B.g. tritici*	I
Isr70	IS	*B.g. tritici*	I
Isr97	IS	*B.g. tritici*	I
Isr204	IS	*B.g. tritici*	V
Isr8	IS	*B.g. tritici*	V
JIW2	UK	*B.g. tritici*	V
WC1110	UK	*B.g. tritici*	V
Ken4-3	US	*B.g. tritici*	V
Isr209	IS	*B.g. tritici2*	A
Isr213	IS	*B.g. tritici2*	I
Isr220	IS	*B.g. tritici2*	I

The isolate name, geographic origin (Switzerland, CH; China, CN; Israel, IS; United Kingdom, UK; United States, US), *forma specialis* (*B.g. tritici,* hexaploid wheat host; *B.g. tritici2*, tetraploid wheat host; *B.g. triticale*, Triticale host), and phenotype (avirulent, A; intermediate, I; virulent, V) on the *pec* mutant are given.

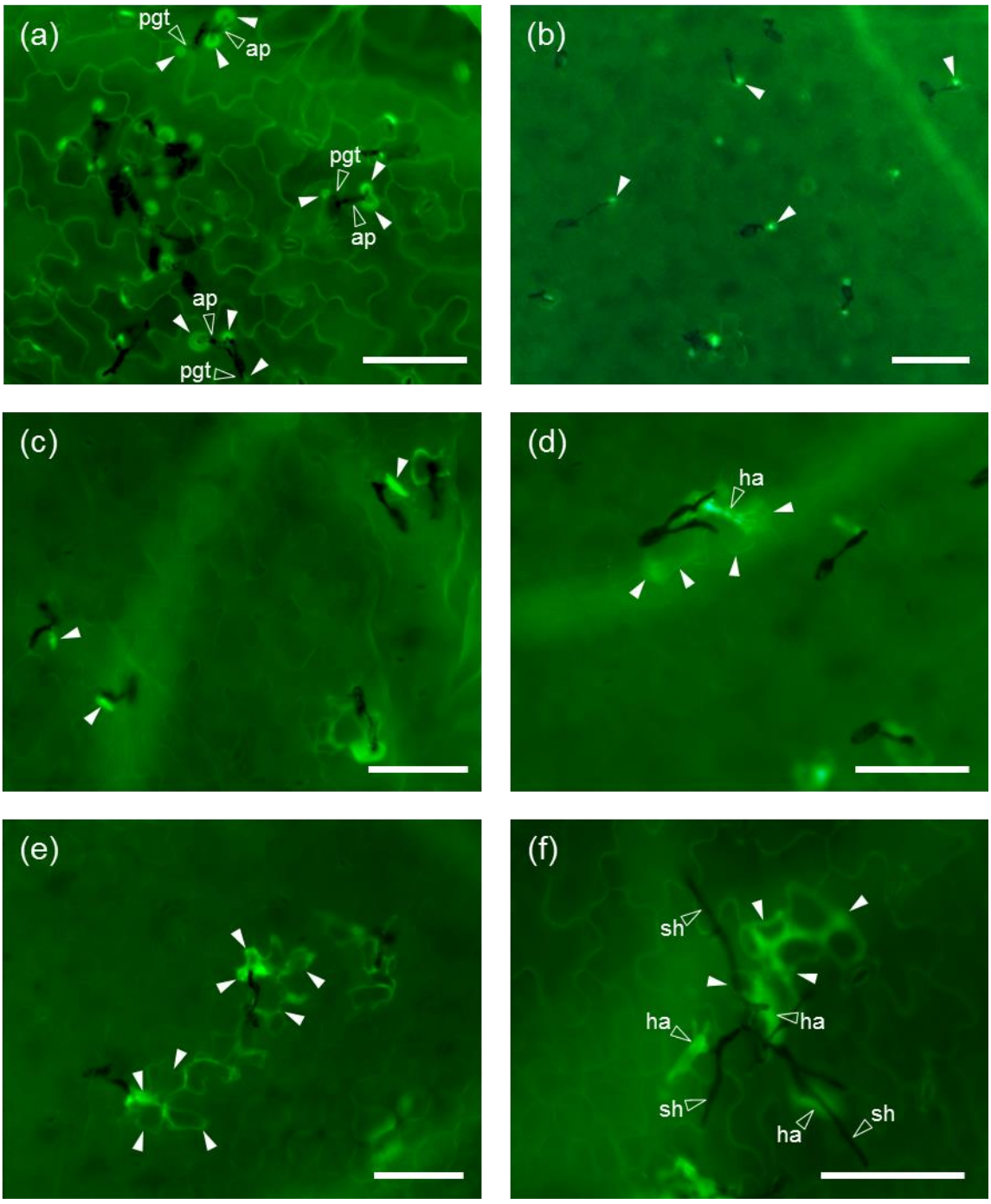


Figure 1. Autofluorescent images of non-host defense responses and visualization of barriers preventing compatible interactions between various *B.g. tritici* isolates and Arabidopsis *pec* mutant. Infected leaf samples were stained with coomassie blue to visualize fungal hyphae and host-derived phenolic defense responses were visualized (see Methods). The scale bar measures 100 µm. (a) Additional aniline blue staining reveals first, second and third contact sites (filled arrows) under the primary germ tube (pgt) and appressoria (ap) of germinated spores. (b) Autofluorescing papillae (filled arrows) prevent attempted penetration by germinated spores. (c) Elongated autofluorescent signal reminiscent of haustorial encasement (filled arrows). (d) Mature haustoria (ha) and induced whole cell HR (filled arrows). (e) The HR visualized as whole-cell autofluorescence (filled arrows). (f) Microcolony with well-established secondary hyphae (sh), multiple haustoria (ha), and whole cell HR (filled arrows).

1E). We visually observed that some isolates were able to occasionally form a mature haustoria, and subsequently secondary hyphae, but it was always followed by cell death before visualization at 7 dpi (Figure 1F). Papillae formation and the phenolics associated with cell death were observed in the majority of interactions, regardless of the isolate tested. Interestingly, isolates which never formed microcolonies often successfully penetrated papillae, but never reached the stage of mature haustorial development due to haustorial encasement.

Virulence in the nonhost interaction is not correlated with geographic origin and does not differ between *B.g. tritici* and *B.g. Triticale* isolates

The 43 geographically diverse isolates were scored 'avirulent' for no microcolony formation, 'intermediate' for occasional, or 'virulent' for often (see methods) forming microcolonies according to microscopic images of infections at 7dpi. The results are depicted in Figure 2A. According to the phenotypes of these 43 isolates, the frequency of race-specific virulence does not differ between isolates from distinct geographic regions. Additionally, despite having an expanded host range, the 8 *B.g. triticale* isolates phenotyped in this study showed no substantial increase in virulence compared to *B.g. tritici* isolates (Figure 2B).

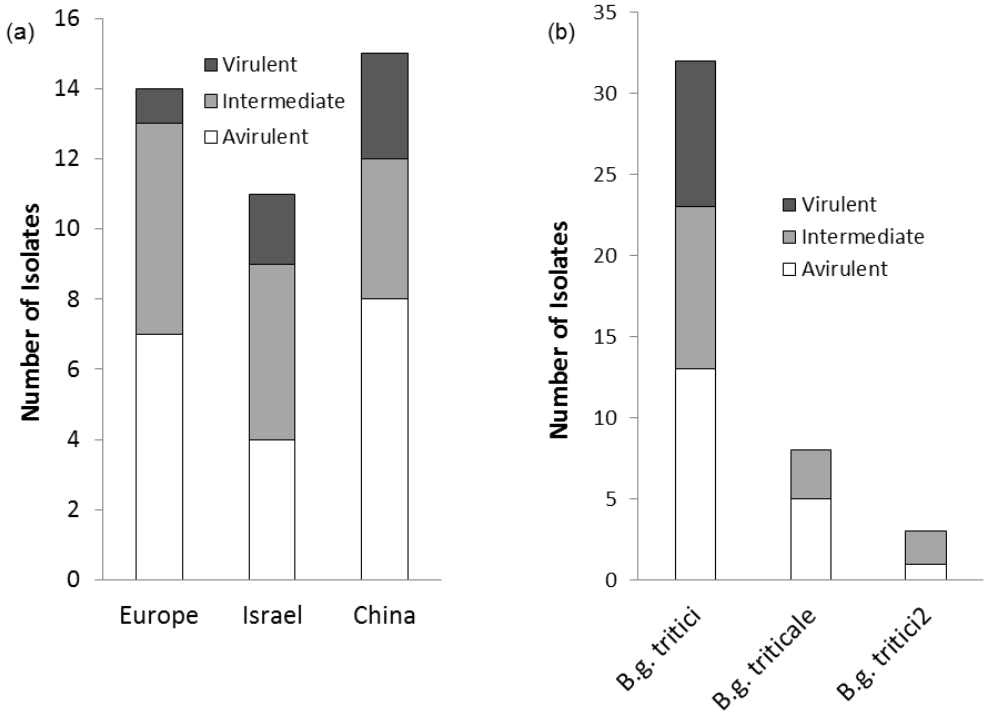


Figure 2. Qualitative phenotypes of 43 *B.g. tritici* and *B.g. Triticale* isolates on the Arabidopsis *pec* mutant. The number of isolates scored as 'avirulent', 'intermediate', and 'virulent' are categorized based on their geographic origin (a) and *formae speciales* (*B.g. tritici,* hexaploid wheat host; *B.g. tritici2*, tetraploid wheat host; *B.g. triticale*, Triticale host) (b).

Variation in the rate of microcolony formation between isolates is significant and quantifiable

In addition to isolates 96224, 96229, 07250, Isr220 and Isr213 that were selected for comparison to previously published quantified virulence levels (Jordan et al, in preparation), 15 additional isolates were assayed for determination of percentage of microcolony formation to obtain a more accurate measurement of virulence differences between isolates. Virulence levels were observed at 7 dpi and microscopic images were counted visually using the ImageJ cell counting software (ImageJ 1.48v). The number of spores which successfully formed microcolonies was compared to the total number of germinated spores (Figure 3). Mean microcolony formation ranged from 0-15% of germinated spores.

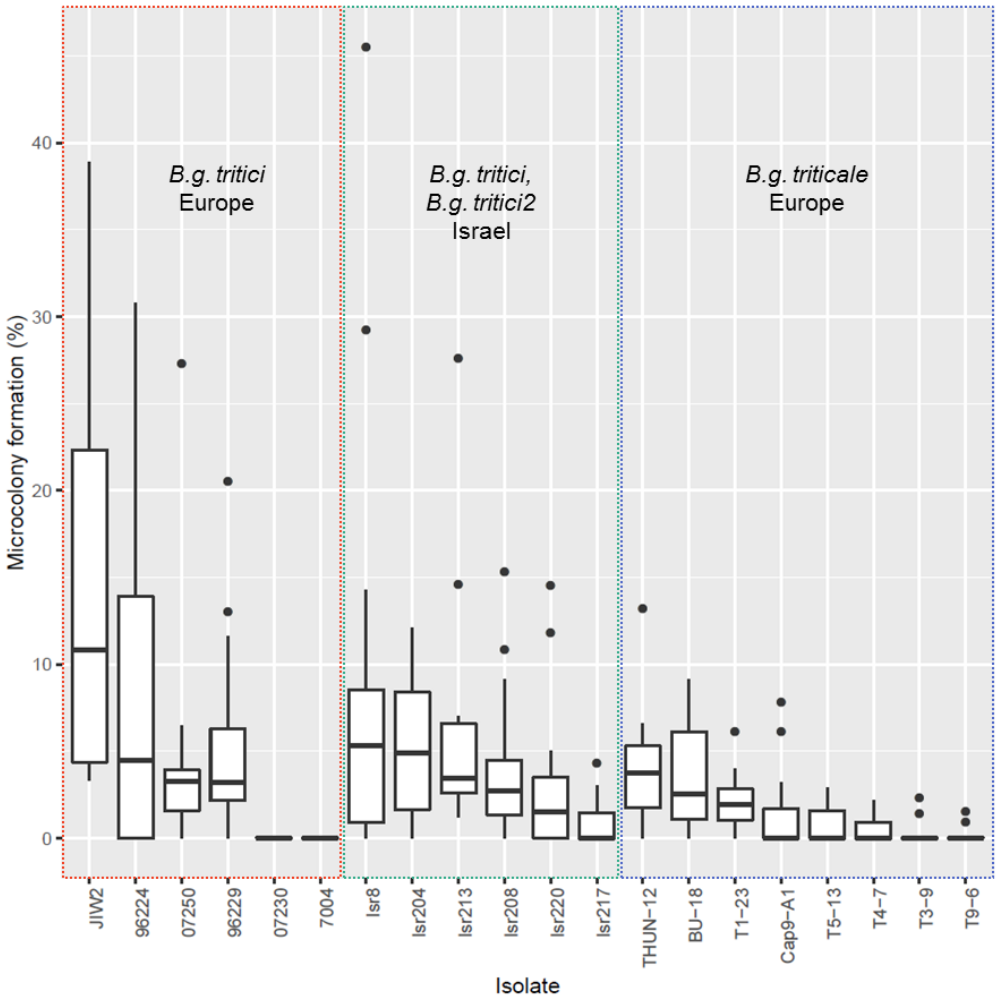


Figure 3. Quantified phenotypes of 20 *B.g. tritici, B.g. tritici2* and *B.g. triticale* isolates. Percent microcolony formation is depicted in a box plot indicating sample quartiles and outliers are shown as black dots. Isolates are divided into European (red box), Israeli (teal box) and *B.g. triticale* (blue box) populations.

Virulence of *B.g. tritici* and *B.g. triticale* isolates on the *pec* mutant varied quantitatively and interestingly, more variation was observed between isolates of *B.g. tritici* than *B.g. triticale*. Among the isolates with quantified phenotypes, two never formed a microcolony (07230, 7004), four were able to occasionally form microcolonies (Isr217, Cap9-A1, T5-13, T4-7), and 14 regularly formed microcolonies (remaining isolates) (Figure 3). This variability could represent the polymorphic presence of either avirulence factors in non-adapted *B.g* ff. spp. that are recognized by the remaining active immune machinery in the *pec* mutant, or virulence factors which aid in colonizing the nonhost species when the major immunity components are compromised. Replicates within isolates varied substantially due to environmental sensitivity of the assay and insufficiently controllable conditions. However, several isolates were either consistently unable (7004, 07230) or able (JIW2) to form microcolonies. This allowed us to select parents for a genetic cross that could be used to map genetic loci responsible for race-specific virulence on a nonhost species.

Preliminary GWAS analysis using binary phenotype assignments

In our first approach to map nonhost virulence factors we performed a preliminary genome-wide association study (GWAS) using 34 of the phenotyped isolates for which we had available complete genome sequences (Table 2). These isolates were assigned binary phenotypes (0, avirulent; 1, intermediate/virulent) to simplify the analysis (16 avirulent, 18 virulent) (Table 2). The binary phenotypes on the *pec* mutants were associated genetically with sequence polymorphisms using the Genome Association and Prediction Integrated Tool (GAPIT) (Lipka et al, 2012) (Figure 4). No SNPs passed the significance threshold, so we identified the 10 best correlated SNPs located in 4 genomic sequence contigs (ctg153, ctg1000, ctg21, and ctg71, Table 3). We checked for genes 10 kb upstream and downstream of each SNP position, resulting in the identification of 10 candidate genes associated with the phenotype of *B.g. tritici*, *B.g. tritici2*, and *B.g. triticale* isolates on the Arabidopsis *pec* mutant (Table 4). It was necessary to re-annotate three genes (Bgt-2048_new, Bgt-2417_new, and Bgt-1154_new). None of the candidates were predicted to contain signal peptides for secretion. SNPs identified by GWAS were found inside the coding sequence of Bgt-2048_new, BgtA-20098, and Bgt-1170, which were annotated as a putative PHD finger and SET domain-containing protein, pyridine nucleotide-disulphide oxidoreductase, and para-aminobenzoate (PABA) synthase genes, respectively (Table 4).

Table 2. Isolates used in the genome-wide association study.

Isolate	*Formae speciales*	Origin	Phenotype
BU-18	*B.g. triticale*	Europe	1
Cap-39	*B.g. triticale*	Europe	1
T1-23	*B.g. triticale*	Europe	1
THUN-12	*B.g. triticale*	Europe	1
9-2	*B.g. tritici*	China	1
3-53	*B.g. tritici*	China	1
6-6	*B.g. tritici*	China	1
1-25	*B.g. tritici*	China	1
46-30	*B.g. tritici*	China	1
36-70	*B.g. tritici*	China	1
6-69	*B.g. tritici*	China	1
96224	*B.g. tritici*	Europe	1
Isr217	*B.g. tritici*	Israel	1
Isr8	*B.g. tritici*	Israel	1
Isr204	*B.g. tritici*	Israel	1
Isr208	*B.g. tritici*	Israel	1
Isr213	*B.g. tritici2*	Israel	1
Isr220	*B.g. tritici2*	Israel	1
T3-9	*B.g. triticale*	Europe	0
T4-7	*B.g. triticale*	Europe	0
T5-13	*B.g. triticale*	Europe	0
T4-6	*B.g. triticale*	Europe	0
10-8	*B.g. tritici*	China	0
15-9	*B.g. tritici*	China	0
5-83	*B.g. tritici*	China	0
14-17	*B.g. tritici*	China	0
12-50	*B.g. tritici*	China	0
13-50	*B.g. tritici*	China	0
11-99	*B.g. tritici*	China	0
41-5	*B.g. tritici*	China	0
Isr15	*B.g. tritici*	Israel	0
Isr209	*B.g. tritici2*	Israel	0
Isr215	*B.g. tritici*	Israel	0
Isr7004	*B.g. tritici*	Israel	0

The isolate name, geographic origin, *formae speciales* (*B.g. tritici,* hexaploid wheat host; *B.g. tritici2*, tetraploid wheat host; *B.g. triticale*, Triticale host), and and binary phenotypes on the Arabidopsis pec mutant used for the GWAS analysis (avirulent, 0; intermediate/virulent, 1) are given.

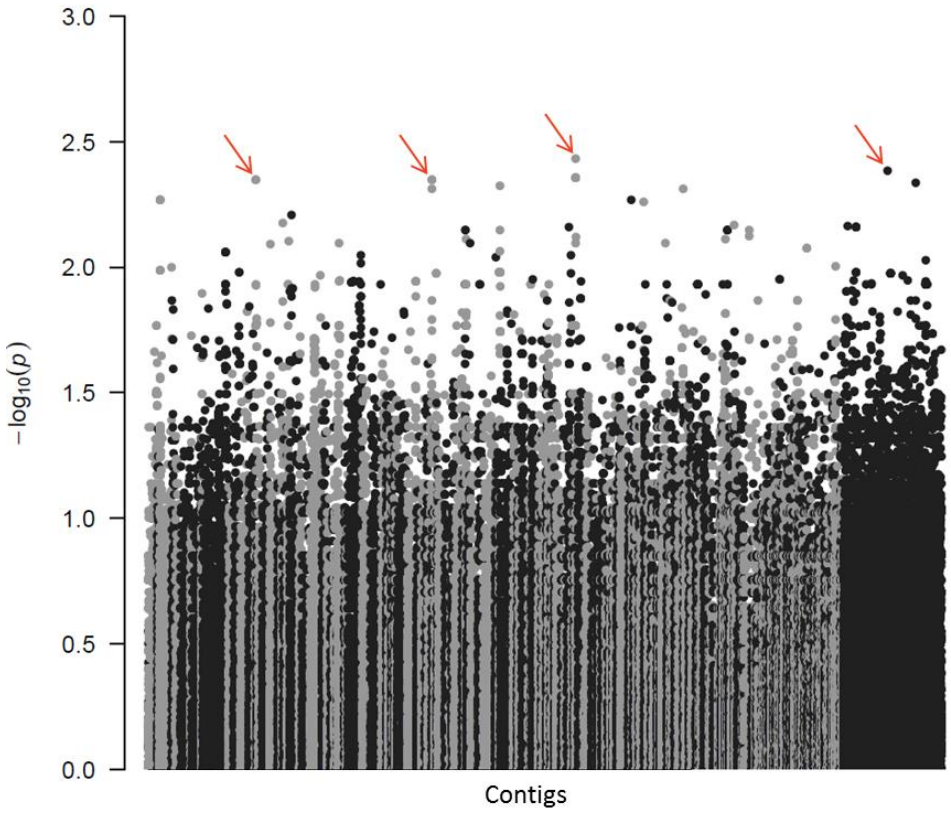


Figure 4. Genome wide association study (GWAS) of the outcome of infection on the Arabidopsis *pec* mutant in a set of 23 *B.g. tritici*, 8 *B.g. triticale*, and 3 *B.g. tritici*2 isolates. Results are depicted at the genome level. Single nucleotide polymorphisms (SNPs) are plotted against the contig position in the genome assembly. Each linkage group is alternately colored black and grey. The genes 10 kb upstream and downstream of the ten best SNPs (located in four contigs: ctg21, ctg71, ctg153, and ctg1000; indicated with red arrows) were selected as candidates and analyzed for polymorphisms.

Table 3. The ten most significant SNPs associated with isolate phenotypes on the Arabidopsis *pec* mutant.

SNP	p-value	Contig	Position
1	0.0037	153	73252
2	0.0041	10000	8842566
3	0.0041	153	93880
4	0.0041	153	77160
5	0.0041	153	78169
6	0.0041	153	94139
7	0.0045	21	157929
8	0.0045	21	164898
9	0.0045	71	814425
10	0.0045	71	814426

The p-value, contig, and position is indicated for each SNP.

Table 4. Candidate genes located within 10 kb of a SNP identified by the Genome-wide association study.

Gene	Contig	Position	Size (AA)	Annotation	Putative Biological Function
Bgt-570	21	174867-175540	198	Regulatory subunit of Nem1p-Spo7p phosphatase holoenzyme	Role in formation of a spherical nucleus and meiotic division
Bgt-2048_new	21	163495-166458	988	Putative PHD finger and SET domain-containing protein	Putative methyl transferase, epigenetic regulator
BgtA-20304	71	821907-823104	359	XAP5 domain containing protein	Circadian clock regulator
BgtA-20098	153	76139-77428	395	Pyridine nucleotide-disulphide oxidoreductase	Disulphide bridge reduction
Bgt-4391	153	84371-84715	115	Hypothetical protein, no signal peptide	None
Bgt-2979	153	85241-86653	384	SYLF domain and C-terminal SH3 domain containing protein	Lipid binding, important for actin polymerization during endocytosis; enzyme regulation, subcellular localization of signalling pathway components, formation of multiprotein complex assemblies
Bgt-2417_new	153	102028-104159	695	Transcriptional regulator CRZ1/similar to C2H2 type zinc finger domain protein	Transcription factor
Bgt-1154_new	153	79771-81358	496	Mannosyltransferase	Call wall/membrace/envelope biogenesis
Bgt-1171	10000	8837948-8838848	262	Alpha 6 subunit of the 20S proteasome	Proteosomal core subunit
Bgt-1173	10000	8839202-8839773	130	Late endosomal protein, probable Vacuolar protein sorting 55 superfamily	Secretion of golgi vacuoles and membrane trakkicking to the vacuole/lysosome
Bgt-1170	10000	8840086-8842614	843	Para-aminobenzoate (PABA) synthase	Chorismate binding, enzyme catalyzes first step of tryptophan biosynthesis

The gene name, contig and position, size, and annotation are given. Putative biological functions associated with the annotations are described.

Establishing a genetic cross for future projects to map both host and nonhost avirulence factors

Given the complexity and quantitative nature of the nonhost virulence phenotype, a genetic cross and QTL mapping is the most straight-forward and reliable strategy for identifying fungal factors involved in nonhost virulence ability. Several isolates showing either complete avirulence or high virulence on the *pec* mutant were screened for compatible mating types. The Swiss isolate 7004 (avirulent on *pec*, MAT1-2) and U.K. isolate JIW2 (virulent on *pec*, MAT1-1) were selected for crossing based on their contrasting phenotypes (Figure 5A-D). JIW2 is the most virulent isolate observed in this study, often forming microcolonies at high frequency, and a larger colony size (Figure 5A), while 7004 was never observed to form a microcolony (Figure 5B). We assessed the HR by visualization of the autofluorescent response, which indicated that JIW2 is more likely to trigger an HR during infection than 7004, which is more often blocked by haustorial encasement (Figure 5C, 5D). The isolates were co-infected onto 3-week old susceptible wheat cultivars and grown for several weeks before allowing the plants to slowly dry out (see Methods). Numerous chasmothecia formed (Figure 5E) and the leaves were stored in an envelope in the dark at room temperature. A survey of the race-specific virulence profiles of the parental isolates revealed that this cross is also polymorphic on several adapted-host (wheat) cultivars (Table 5). Therefore, in addition to be polymorphic for virulence in a nonhost interaction, this cross could also be used to genetically map *AvrPm3b, AvrPm3d, AvrPm3e, AvrPm4b, AvrPm5b, AvrPm13, AvrPm16,* and *AvrPm17*.

Parental isolates JIW2 and 7004 are polymorphic for key melanin biosynthesis genes

Although powdery mildews are generally considered colorless, the parents of the genetic cross, JIW2 and 7004, appear to produce different amounts of melanin. When inoculated onto detached leaf segments, 7004 conidiophores are typically more 'white', while JIW2 is more 'pink' or melanized. It is well known that melanization plays an important role in virulence, particularly for appressorial penetration. We used BLAST to identify the genes related to melanization from *B.g. tritici* and identified 14 candidates (Table 6). We compared the sequences of these genes between 7004, JIW2, and the reference isolate 96224, which has a melanization phenotype similar to JIW2. Several genes from JIW2 were partial or truncated likely because of low sequence quality of the JIW2 genome sequence. This prevented direct comparison between the two parents. Ten of the 14 genes were polymorphic between isolates (Figure 6).

Table 5. Phenotypes of the two parental isolates, JIW2 and 7004 on *Pm* resistance gene containing wheat cultivars.

Cultivar/line	Gene/allele	Source	7004	JIW2
Axminster/8*CC	*Pm1a*	*T. aestivum*	A	A
M1N	*Pm1c*	*T. aestivum*	V	V
Weihenstephan I	*Pm1c*	*T. aestivum*	V	V
C12632	*Pm2*	*T. aestivum*	V	V
Ulka/8*CC	*Pm2*	*T. aestivum*	V	V
Asosan8*	*Pm3a*	*T. aestivum*	V	A
Chul8*	*Pm3b*	*T. aestivum*	V	A
Sonora8*	*Pm3c*	*T. aestivum*	V	V
Kolibri	*Pm3d*	*T. aestivum*	I	A
W150	*Pm3e*	*T. aestivum*	I	A
M.Amber	*Pm3f*	*T. aestivum*	V	V
Khapli/8*CC	*Pm4a*	*T. dicoccum*	V	V
Ronos	*Pm4b*	*T. carthlicum*	I	V
Hope/8*CC	*Pm5a*	*T. dicoccum*	V	V
Kormor0n	*Pm5b*	*T. aestivum*	V	A
I5	*Pm5d*	*T. aestivum*	V	V
14/2*Starke	*Pm6*	*T. timopheevii*	V	V
N14	*Pm9*	*T. aestivum*	V	V
Wembley	*Pm12*	*Ae. Speltoides*	A	A
Pm13	*Pm13*	*T. longissimum*	A	V
Pm16	*Pm16*	*T. dicoccoides*	A	V
T0m107	*Pm17*	*Secale cereale*	A	V
Amigo	*Pm17*	*Secale cereale*	A	V
Pm19	*Pm19*	*Aegilops tauschii*	V	V
T0m W-104	*Pm20*	*Secale cereale*	V	V
Pm24	*Pm24*	*T. aestivum*	V	V
Pm29	*Pm29*	*T. aestivum*	V	V
Pm32	*Pm32*	*Ae. Speltoides*	V	V
5-BIL29 (durum)	*Pm36*	*T. dicoccoides*	A	A
Pm43	*Pm43*	*Thinopyrum intermedium*	V	V
Inbar	*Susceptible*	*T. durum*	V	V

The name of the wheat cultivar is given, and the *Pm* gene it carries, as well as the source of the gene and whether the isolates are avirulent (A), intermediate (I), or virulent (V) on the cultivar.

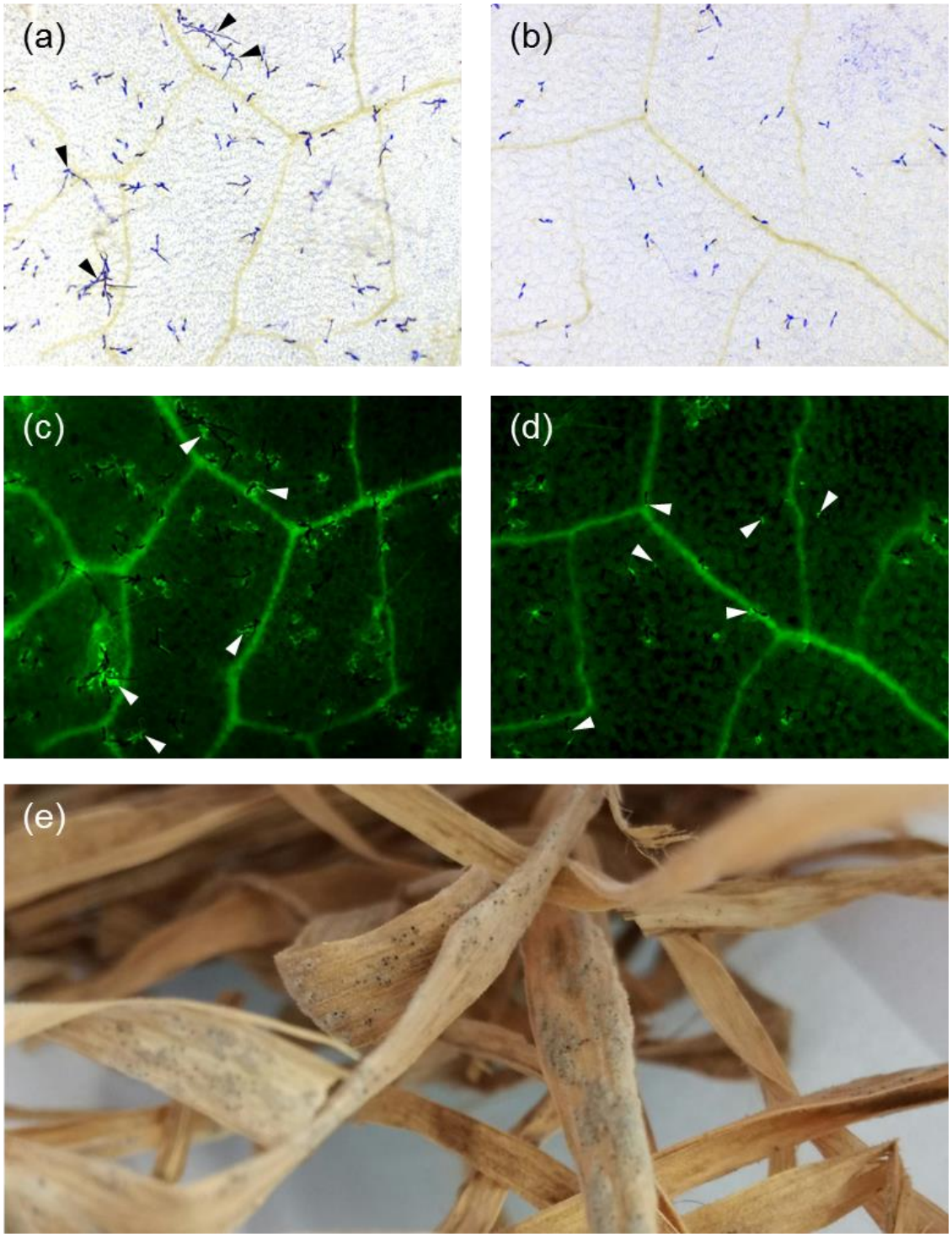


Figure 5. The genetic cross of JIW2 and 7004. Coomassie blue staining revealed microscopic differences in the frequency of microcolony formation (black arrows) between the virulent parent, JIW2 (A), and the avirulent parent, 7004 (B). Autofluorensce in response to the virulent isolate, JIW2, includes whole cell death (white arrows) (C), while papillae and haustaurial encasement (white arrows) prevent infection by the avirulent isolate, 7004 (D). Dense chasmothecia formed by crossing JIW2 and 7004 (E).

Table 6. Predicted melanin biosynthesis-related genes in the *Blumeria graminis* f. sp. *tritici* genome.

Database reference	Bgt Gene	Annotation	%id	alignment length	e-value	bit score	SNPs: 96224 vs. 7004	non-syn	SNPs: 96224 vs. JIW2	JIW2 annotation	Notes about annotations
gi_296239577_gb_ADH01674.1	Bgt-5_p	Polyketide synthase	41.34	1761	0.0E+00	1301	7	5	4	partial	
gi_1026736885_gb_AND66116.1	BgtA-21367_p	Tyrosinase P	22.44	303	9.0E-04	39.3	5	5	2	truncated	no stop codons, messy end
gi_7141324_gb_AAF37291.1	Bgt-834_p	Transcription factor	45.61	57	1.0E-07	53.5	3	4	2	truncated	
gi_12862766_dbj_BAB32575.1	Bgt-1212_p	Conidial pigment biosynthesis oxidase (Laccase)	40.65	556	5.0E-114	406	2	2	2	truncated	no stop codons
gi_339473128_gb_EGP88222.1	BgtASP-20119_p	Tyrosinase precursor (Monophenol monooxygenase)	29.48	597	9.0E-72	266	2	2	3	truncated	no stop codons
gi_1042851558_sp_E9QUT3_1	Bgt-1153_p	Conidial pigment biosynthesis oxidase (Hydroxynaphthalene reductase)	30.81	185	2.0E-16	80.9	4	1	0	partial	
gi_3024596_sp_P56221.1	Bgt-5146_p	Scytalone dehydratase	21.67	60	2.6E+00	26.2	2	1	1	truncated	
gi_1042851631_sp_Q4WZB4_2	Bgt-1213_p	Conidial pigment biosynthesis oxidase (Multicopper oxidase)	41.61	560	3.0E-120	427	0	1	1	full	
gi_1008818903_sp_F9X365_1	Bgt-5379-2_p	Developmental and secondary metabolism regulator (Velvet complex subunit 1)	55.8	224	3.0E-60	226	1	1	1	partial	no stop codons
gi_74671809_sp_Q4WQL4.1	Bgt-596_p	Developmental regulatory protein	55.93	59	2.0E-12	68.6	1	1	2	full	
gi_13570027_gb_AAG29497.2	Bgt-1160_p	Tetrahydroxynaphthalene reductase	22.71	273	4.0E-13	69.7	1	0	2	full	
gi_1026736883_gb_AND66115.1	Bgt-1140_p	Aspulvinone E synthetase	28.91	460	5.0E-28	121	0	0	0	truncated	
gi_74674843_sp_Q4WZB3_1	Bgt-196_p	Conidial pigment biosynthesis protein (Heptaketide hydrolyase)	23.17	164	9.6E-02	32.7	0	0	0	full	
gi_4520324_dbj_BAA75888.1	Bgt-3573_p	Transcription factor	38.82	85	4.0E-11	65.5	0	0	0	truncated	

The NCBI database reference ID is indicated, followed by the gene name and annotation from the *B.g. tritici* genome. The total number of single nucleotide polymorphisms (SNPs) in each parent sequence are enumerated, as well as the number of nonsynonymous residue changes. The status of the JIW2 annotation for each gene is either full, partial, or truncated.

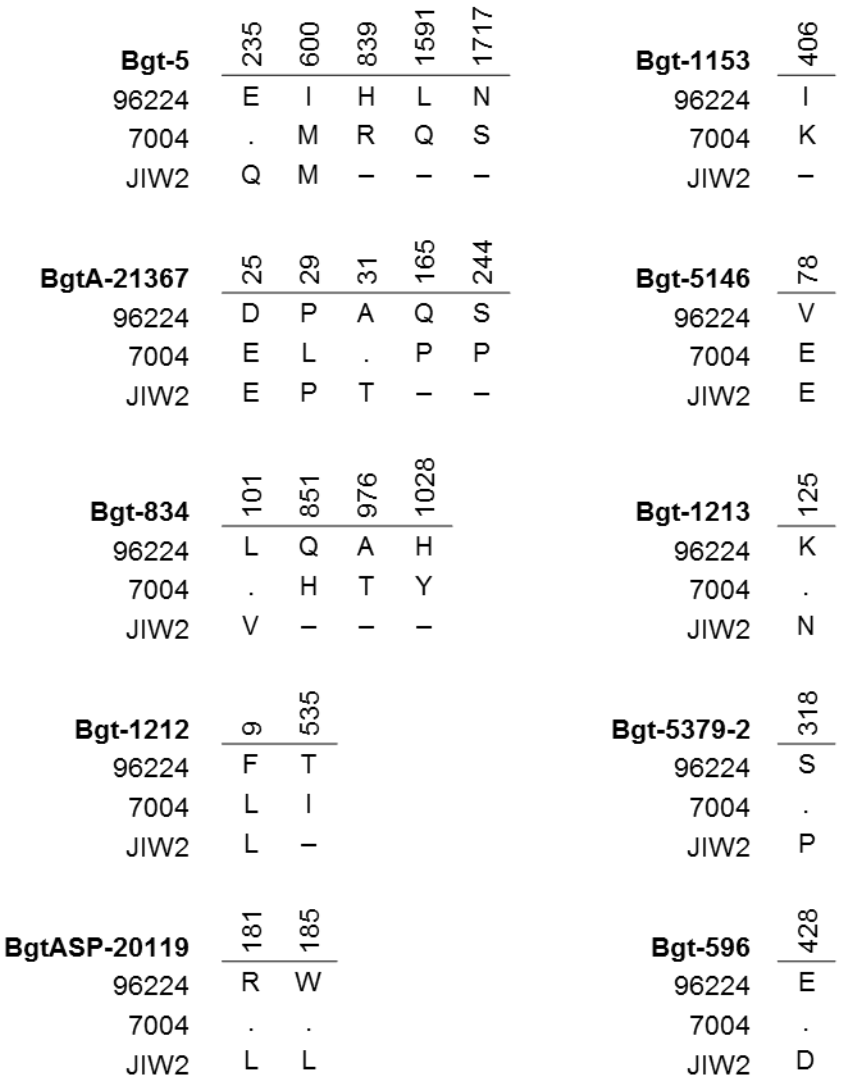


Figure 6. Nonsynonymous residue changes in melanin biosynthesis related genes between JIW2, 7004, and the reference isolate 96224. Amino acid polymorphisms compared to the reference isolate are depicted using the one-letter code for amino acids. Dashes indicate missing data and dots indicate identical residues.

We focused on polymorphisms in Bgt-5, the polyketide synthase (PKS) ortholog, since it was previously implicated in melanization in the hemibiotroph *Zymoseptoria tritici* (Lendenmann et al, 2014). Bgt-5 contains several polymorphisms between the tested *B.g. tritici* isolates, including 5 nonsynoymous substitutions (residues 600, 839, 1591, and 1717) between 7004 and the reference isolate 96624 (Figure 6). Four of the nonsynonymous substitutions occur in functional domains of the PKS protein (Figure 7). Alterations to these domains may result in differential melanin production, and these polymorphisms should be tested for co-segregation with the phenotype once the mapping population is available.

Figure 7. Polymorphisms in the Polyketide Synthase (PKS) gene. The full length PKS gene and functional domains are depicted to scale. Synonymous nucleotide substitutions are indicated by black vertical dashes, while nonsynonymous substitutions are indicated by red vertical dashes.

Recent evidence suggests that a transcription factor located near to the PKS in *Z. tritici* might cause the differential melanization phenotype (McDonald, personal communication). We checked whether the transcription factor (Bgt-2417_new) identified by GWAS was located near to the PKS (Bgt-5) in *B.g. tritici*. Bgt-2417_new is located in Bgt_pacbio-9, while Bgt-5 is located in Bgt_pacbio-10, two unlinked pacbio scaffolds.

2.3 Discussion

Nonhost resistance is generally considered to involve multiple molecular interactions between pathogen and non-adapted host, in addition to preformed immunity barriers (Senthil-Kumar and Mysore, 2013). Here, we studied the interaction between *Blumeria graminis* f. sp. *tritici* and a non-adapted host, *Arabidopsis*, with the goal of identifying factors that confer race-specific virulence. We screened genetically diverse powdery mildew isolates of wheat and triticale on the immune deficient *Arabidopsis pec* mutant and concluded that differences in virulence are not associated with the geographic origin of isolates and do not differ between *B.g. tritici* and *B.g. triticale*. However, we observed race-specific arrest of the progression of infection, either by papillae formation, haustorial encasement, or cell death. Differential disease progression has been previously observed in this system, but only between different pathogen species (*B.g. hordei, Erysiphe pisi*, and *Golovinomyces orontii*) and not between isolates (Lipka et al, 2005). Around 90% of the *B.g. tritici* and *B.g. triticale* isolates tested were able to form occasional microcolonies, up to an average frequency of 12% of germinated spores, while for 10% of isolates the formation of a microcolony was never observed. This finding is consistent with previous reports of race-specific virulence of *B.g. tritici* on the *pec* mutant (Jordan et al, in preparation), but has not been reported for other *formae speciales*, particularly the extensively studied *B.g. hordei* from barley. We suggest that polymorphic factors on the pathogen side determine the most frequent stage at which infection is arrested in this interaction.

Distinguishing virulence and avirulence factors in haploid pathogens is difficult or impossible

What is not clear at the onset of such a study is whether the identified factors governing this nonhost interaction confer virulence or avirulence. Because *Blumeria* is a haploid organism, it can be difficult or impossible to ascertain from genetics whether a factor is dominant or recessive, and whether its function is activity or inactivity. In the case of the more avirulent *B.g. tritici* isolates, it could be that they harbor effectors which elicit ETI in the *Arabidopsis* nonhost. It could also be that some versions of effectors are more or less able to fulfill their virulence functions on the nonhost machinery. In this case, an isolate harboring such an effector would be better able to manipulate the nonhost and therefore be more virulent.

Mechanistic differences between *pec* and *pps* may contribute to differing virulences of Blumeria ff. spp.

Previously Jordan and colleagues (in preparation) demonstrated that the *pec* mutant allows race-specific microcolony formation by isolates of *B.g. tritici* (5-35%), while the *pps* mutant consistently allowed low levels of microcolony formation (10%). It is unclear what the differences in these mutant backgrounds are which might explain the variation in non-adapted pathogen colonization. The *pps* mutant lacks functional PEN2, PAD4, and SAG101, while the *pec* mutants lacks functional PEN1, EDS1, and contains a mutation in the CAMTA3 transcription factor (*camta3-4d*), which constitutively silences several defense-related genes (Galon et al, 2008). Isolates of *B.g. hordei* are equally capable of forming microcolonies on the *pps* mutant as the *pec* mutant. This suggests that all tested *B.g. hordei* isolates encode a factor that gives the *formae speciales* an advantage in both immune-deficient mutant backgrounds.

Enhanced colonization of the *pec* mutant when compared to the *pps* mutant suggests that *B.g. tritici* is more readily able to penetrate in the absence of a functional *PEN1* rather than a nonfunctional *PEN2*. *PEN1* forms a SNARE complex with SNAP33, VAMP721/VAMP722 and is involved in but not required for trafficking of exosomes carrying potentially toxic compounds to the site of penetration (Kwon et al, 2008; Pajonk et al, 2008; Nielsen et al, 2012). Restriction of invasive growth by *PEN1* has only been demonstrated for non-adapted powdery mildews and not for other microbial pathogens (Collins et al, 2003, Lipka et al, 2005). It was hypothesized that this is due to the ability of adapted mildews to detoxify *PEN1* secreted compounds (Meyer et al, 2009). This specificity of powdery mildews to elicit an accumulation of *PEN1* is also observed in barley, where the *PEN1* syntaxin homolog *ROR2* also specifically accumulates under appressorial contact sites (Bhat et al, 2005), and is required for full *mlo* mediated resistance (Consonni et al, 2006). Although it is still unclear what role

PEN1 plays in resistance, it would be interesting to see whether a homolog is also conserved in wheat, an absence or alteration of which might explain the differences in the effectiveness of this pathway in *Arabidopsis* against host-specialized *formae speciales*.

A second possible explanation is that the different mutated EDS1/SAG101/PAD4 signaling components, which together form an NLR signaling complex but that have been shown to have distinct alternative functions (Rietz et al, 2011), might have complex interactions with *B.g. tritici* effectors. However it is not clear why this interaction would be different for *B.g. tritici* and *B.g. hordei*. Given that wheat and barley are so closely related, the cellular signaling machinery in the hosts is likely to be similar. Regardless, if differences in the host signaling responses contribute significantly to the virulence phenotype, it is likely to be indirectly caused by effector polymorphisms in the pathogen. A third possibility is the relatively unknown role of the *camta3-4D* transcription factor, which silences several defense-related genes (Galon et al, 2008). It would require further investigation into the transcriptional effects of this gene to hypothesize about its role in the interaction.

Overall, with the correct combination of mutated host-immune machinery (*pen1/eds1/camta3-4D*), *B.g. tritici* is more able to overcome the first level of basal resistance that would normally prevent penetration and haustorial development, and interestingly, at the next stage the interaction becomes quantitative. What factors could be at play in determining this quantitative virulence phenotype? No significant differences in the metabolic machinery or morphology between *B.g. tritici* isolates have been described. Therefore, it is likely that specific components of the effector repertoire in the virulent parent are either more adept at manipulating the *Arabidopsis* nonhost machinery, or effectors in the avirulent isolate are recognized by nonhost NLRs. Most likely it is a combination of these, and is likely to be strongly influenced by any polymorphism in melanin production. The contribution of each of these factors will become clearer following a QTL analysis.

Virulence on the nonhost *Arabidopsis* is likely to be a quantitative trait

Thus far the evidence from nonhost studies suggests a major role for effectors in deciding the outcome of a nonhost interaction, both by effector recognition by plant receptors and the effects of host-specific effector function (Stam et al, 2014). Host-specific effector function is the result of effector evolution to manipulate targets in the adapted hosts, resulting in a divergent effector which may not be able to manipulate a non-adapted host target. Loss of effector virulence function would limit pathogen virulence and may result in stunted growth or an inability to effectively reproduce.

The same is true if the pathogen cannot successfully take up essential nutrients from a nonhost plant, particularly in the case of obligate biotrophic pathogens, who depend entirely upon the host for growth and reproduction. Obligate pathogens can lose the machinery to metabolize essential nutrients and therefore depend upon successful production and uptake from the host (Spanu et al, 2010; Baxter et al, 2010). Several genes involved in inorganic nitrogen (nitrate) assimilation are lacking in rust fungi, specifically *Melampsora larici-populina* and *Puccinia graminis* f. sp. *tritici* (Slot and Hibbett, 2007). In *Blumeria graminis* several genes involved in primary metabolism are lacking, notably those involved in thiamine biosynthesis, anaerobic fermentation, biosynthesis of glycerol from glycolytic intermediates, and, as with the rust fungi, genes involved nitrate assimilation (Spanu et al, 2010). Small differences in metabolic machinery on either the pathogen or host might have large effects during nonhost interactions. Taken together, it is likely that a QTL mapping approach would yield many minor loci contributing towards virulence (or avirulence) of the non-adapted pathogen on the *Arabidopsis pec* mutant.

Virulence on the *Arabidopsis pec* triple mutant differs between *formae speciales*

Despite being a close relative, *B.g. hordei* (barley powdery mildew) shows increased virulence on the *pec* mutant compared to *B.g. tritici* according to Jordan and colleagues (in preparation). However, it is not likely that significant genetic differences account for variation in virulence between *formae speciales* since the phylogenetic distance between these close relatives is much less than that of their adapted hosts and the nonhost *Arabidopsis*. This suggests that the causative genetic factors of race-specific virulence in *B.g. tritici* might not be that numerous. We can also infer that PAMP recognition is not likely significantly involved, as these are most assuredly shared by the *formae speciales.*

Another close relative that we tested, the previously described *B.g. triticale* isolates, is known to be the result of a hybridization of *B.g. tritici* and *B.g. secalis* (rye powdery mildew) (Menardo et al, 2016) and in addition to wheat, gained the ability to infect a 'new' mildew host, Triticale (an amphiploid of wheat and rye). Given their virulence on wheat and recent host jump onto Triticale, there might be factors in the genetic background *B.g. triticale* similar to those of *B.g. hordei* that confer an advantage when infecting nonhost species. Interestingly, we did not observe any *B.g. triticale* isolates with high enough rates of microcolony formation to be considered as virulent. More sampling, including *B.g. secalis* isolates, will clarify whether this is a true distinction between the *formae speciales*, and resolve whether this would be an interesting topic of further study.

Additional isolates and greater phenotype precision are necessary for significant GWAS results

With a revised version of the *B.g. tritici* genome pending and an extensive international *B. graminis* genome collection, with more sampling it should be possible to use a Genome-Wide Association Studies (GWAS) to determine whether any QTLs of major effect control the interaction between *B.g. tritici, B.g. triticale,* and the *pec* mutant. Many isolates with sequenced genomes are available in our collection and therefore suitable for a GWAS analysis. A GWAS using the quantitative phenotypes of 34 isolates for which complete genome sequences were available was not sufficient to yield a significant result, although in another study genetic association was successful in identifying candidate avirulence factors when applied to only 17 *B.g. hordei* isolates (Lu et al, 2016). GWA studies to identify single causative factors of a phenotype are much more robust than those that attempt to associate multiple factors to one phenotype. Therefore, pursuing a GWAS study with qualitative data will require many more isolate phenotypes. Alternatively, using quantitative data would require less isolates for a more robust result.

The phenotyping assay is limited by low virulence and the instability of phenotypes

Stability of the phenotyping assay is a limiting factor for studies of specificity. The values for microcolony formation in this study are lower than those obtained by Jordan and colleagues (in preparation), and low levels of virulence are problematic. The assay is highly sensitive to known and unknown environmental conditions, which results in high variability within and among replicates. Different sections of the infected leaf are variably prone to allowing microcolony formation. For example, often large extended colonies were found on the leaf stem, and higher microcolony frequencies were observed on the leaf apex compared to the rest of the leaf. Since only parts of the leaves can be assayed using the image-based quantification, this introduces a bias in the sampling. Fortunately, the presence of completely avirulent isolates (7004) greatly improves the power of phenotyping in a mapping population. By maximizing virulent and avirulent phenotypic distinctions, finding significant differences in the progeny is more likely. The complete avirulence of the 7004 isolate might be due to a dominant avirulence factor, which could be easily identified in the mapping population. However if the phenotype differs quantitatively within the population and no clear segregation pattern is observed, accurate progeny phenotypes will be crucial to compare virulence profiles between isolates when the rate of microcolony formation is so low. Accordingly, a high number of replications and precise quantification is a more robust and precise, albeit more time-consuming, method, but will be necessary for a powerful QTL analysis.

Differences in melanization potentially resulting from polymorphisms in the polyketide synthase gene may explain differences in penetration efficiency between JIW2 and 7004

Other challenges that non-adapted pathogens face which may limit their success in nonhost interactions are the inherited differences in physiology and toxic secondary metabolites when infecting a nonhost plant. The frontline against pathogen attack is the host cell wall, which can differ in its callose, lignin and wax content (Malinovsky et al, 2014). Many plant pathogens employ melanin for generating sufficient turgor pressure to pierce through cell walls, as well as a redox buffer against free radicals (Howard and Valent, 1996; Jacobson, 2000). When QTL mapping melanization, fungicide and cold tolerance in the wheat fungal pathogen *Zymoseptoria tritici*, Lendenmann and colleagues (2014) identified polymorphisms in the polyketide synthase (PKS) enzyme, the gene which catalyzes the first step of DHN melanin biosynthesis. These polymorphisms make the PKS gene the most probable candidate for the large effect QTL, explaining over 45% of the variance for fungicide tolerance. Given that we see in our collection that melanin content can vary macroscopically between *B.g. tritici* isolates, and that it can be crucial for penetration ability and is an important anti-toxin, it is plausible that polymorphisms in the melanin metabolic machinery could contribute toward the virulence trait. The two parents of the genetic cross, JIW2 and 7004, visibly differ in their melanin production. JIW2 is more 'pink', or melanized, particularly at later stages of development, and 7004 is predominantly white or colorless until very late in conidiophore production. We compared the sequences of our 14 candidate genes (Table 4), however direct comparison between 7004 and JIW2 was difficult due to the low sequence quality of the JIW2 genome. We identified 5 nonsynonymous substitutions between 7004 of the 96224 reference isolate that occur in functional domains of the PKS protein (Figure 7). The substitution at position 600 is unlikely to be the causative substitution because it is shared between 7004 and JIW2 according to the partial PKS sequence from JIW2. The substitution at position 839 does not significantly change the polarity, acidity or hydropathy of the residue. Therefore the L1591Q and N1717S substitutions are the best candidates from the PKS gene to explain the difference in melanization between 7004 and JIW2.

2.4 Materials and Methods

Plant Material

To make the *pec* mutant, *Col-0 camta3-4D pen1* mutant was crossed with the *eds1-2* mutant, in which the *Ler*-derived *eds1-2* allele had been introgressed into the Col-0 background (Bartsch et al, 2006). *Arabidopsis* seeds were surface sterilized with 1% sodium hypochlorite, 0.03% Triton X-100, stratified for 3d at 4°C, and plated on half-

strength Murashige and Skoog medium containing 0.6% phytagel, 2% sucrose. Plates were incubated at a 16-h-light/8-h-dark cycle at 22°C. For infection tests 10 day old seedlings were transferred to soil and grown in a 16-h-light/8-h-dark cycle at 22°C.

Fungal isolates and Arabidopsis *infection assays*

Blumeria graminis isolates (*B.g. tritici* and *B.g. triticale*) were collected from around the world and are currently maintained in our own isolate collection. Isolates were propagated on the susceptible wheat cv. Kanzler as described by Parlange et al. (2011). For all experiments, plant material or detached leaves with or without fungal spores were kept at 20°C, 70% relative humidity and 16 h of light. Three to four week-old *Arabidopsis* rosette leaves were harvested and placed on 0.8% water agar plates supplemented with 0.03 g/l benzimidazole. To inoculate, fungal isolates were dusted over plates. After incubation for 1 h in darkness, plates were incubated under short day conditions in a 8-h-light/16-h-dark cycle at 22°C in a climate controlled growth chamber.

Microscopic evaluation and image analysis

Seven days after inoculation detached leaves were cleared using chloral hydrate (100 g in 40 mL H_2O) and stained with coomassie blue to visualize fungal structures. Callose deposition at the contact sites were stained using aniline blue. Leaves were examined by light microscopy and imaged using a digital color camera and 20x objective (Axiocam HR, Carl Zeiss AG, Germany). Autofluorescence was visualized using excitation in the blue-light range, from 450-490 nm and the emission recorded from 500-550 nm. At least four leaves from different plants were inoculated per isolate, and four images from each leaf were assessed qualitatively. Qualitative evaluation was based on the frequency of the presence of at least one microcolony in an image. For each isolate, a total of 16 images (4 x 4 leaves) were assessed for no (0 pictures), occasional (1-8 pictures) and often (>8 pictures) microcolony formation. Quantification was done by counting at least 100 spore-cell interactions per leaf using the ImageJ cell counter plugin (Rasband, W.S., ImageJ, U. S. National Institutes of Health, Bethesda, Maryland, USA, http://imagej.nih.gov/ij/, 1997-2016).

Genetic cross of two B.g. tritici *isolates*

The cross between the Swiss isolate 7004 (MAT1-2) and the UK isolate JIW2 (MAT1-1) was performed as described by Wicker et al. (2013). Successful chasmothecia formation was contingent upon sufficient day-time temperature, air circulation, and slow reduction in water availability during the host drying phase. Crosses were performed in the greenhouse in March/April 2016.

CHAPTER 3

Multiple avirulence loci and allele-specific effector recognition control the *Pm3* race-specific resistance of wheat to powdery mildew

Salim Bourras*[1], Kaitlin Elyse McNally*[1], Roi Ben-David[1,2], Francis Parlange[1], Stefan Roffler[1], Coraline Rosalie Praz[1], Simone Oberhaensli[1], Fabrizio Menardo[1], Daniel Stirnweis[1,3], Zeev Frenkel[4], Luisa Katharina Schaefer[1], Simon Flückiger[1], Georges Treier[1], Gerhard Herren[1], Abraham B. Korol[4], Thomas Wicker[1] and Beat Keller[1].

*These authors contributed equally to this work.

[1]Institute of Plant Biology, University of Zurich, Zollikerstrasse 107, CH-8008 Zürich, Switzerland;
[2] Current address: Institute of Plant Sciences, Agricultural Research Organization, Volcani Center, 50250 Bet Dagan, Israel;
[3] Current address: KWS Saat SE, Grimsehlstrasse 31, 37555 Einbeck, Germany;
[4] Institute of Evolution, University of Haifa, Mount Carmel, 31905 Haifa, Israel.

Published in The Plant Cell, Issue 27, page 2991-3012, October 2015

Supplemental Figures, Tables, and Text in Appendix I

Abstract

In cereals, several mildew resistance genes occur as large allelic series. In wheat, 17 functional *Pm3* alleles confer agronomically important race-specific resistance. While the molecular basis of race-specificity is characterized in wheat, little is known about the corresponding avirulence genes in powdery mildew since no avirulence gene for *Pm3* has been cloned. Here, we dissected the genetics of avirulence towards six *Pm3* alleles and found that it is controlled by three major *Avr* loci, with a common *locus_1* involved in all *AvrPm3/Pm3* interactions. We cloned the effector gene *AvrPm3*$^{a2/f2}$ from *locus_2* which is recognized by the *Pm3a* and *Pm3f* alleles. Specificity was demonstrated by induction of a *Pm3* allele-dependent hypersensitive response in transient assays in *Nicotiana benthamiana* and in wheat. Gene expression analysis of *locus_1* encoded gene *Bcg1* and *AvrPm3* $^{a2/f2}$ revealed significant quantitative differences between isolates, indicating that in addition to protein polymorphisms, expression levels play a role in avirulence. We propose a model involving three components for race-specificity: an allele-specific avirulence effector, a resistance gene allele, and a pathogen-encoded suppressor of avirulence. Thus, whereas specificity is controlled by a genetically simple allelic series in wheat, recognition on the pathogen side is more complex, allowing flexible evolutionary responses and adaptation to resistance genes.

3.1 Introduction

The gene-for-gene interaction is a genetic model of plant-pathogen interactions, wherein a single plant disease resistance gene (*R*) and a single complementary avirulence gene in the pathogen (*Avr*), account for pathogen recognition resulting in *Avr-R* mediated resistance (Flor, 1971; Jones and Dangl, 2006). Specific recognition of one of the many pathogen secreted virulence factors, also known as effectors (Rafiqi et al., 2012), leads to *Avr-R* coevolution, which imposes strong reciprocal selection leading to an evolutionary arms race, where pathogens adapt to avoid recognition by plant *R* genes, and plants counter-adapt to regain recognition of pathogen effector genes (Jones and Dangl, 2006; Boller and He, 2009). In agreement with the gene-for-gene model, all of the race-specific R proteins mediate resistance upon recognition of their cognate AVR factors. *R* genes typically encode intracellular proteins with an N-terminal coiled-coil (CC) or TOLL/interleukin-1 receptor (TIR), a central nucleotide-binding site (NBS) and a C-terminal leucine-rich-repeat (LRR) domain (Marone et al., 2013), and many induce hypersensitive response (HR), a form of programmed cell death that is effective particularly against biotrophic pathogens (Moffet et al., 2002).

Studies of multi-allelic race-specific *R* genes suggest that they have evolved under strong diversifying selection to recognize specific *Avr* alleles from the pathogen (Rose et al., 2004; Bhullar et al., 2009; Kanzaki et al., 2012). Interestingly, there are only few examples

of true allelic series of race-specific *R* genes in plants which occur as a single gene locus with distinct allelic specificities, and differing spectra of pathogen races recognized (Ellis et al., 1999; Rose et al., 2004; Seeholzer et al., 2010; Kanzaki et al., 2012). Their evolutionary history contrasts with the more frequent resistance genes which are members of larger gene families, organized in complex clusters of gene paralogs, and which can evolve through unequal recombination as shown for example for the flax *M* locus (Anderson et al., 1997), the tomato *Cf4/9* locus (Parniske et al., 1997) and the maize *Rp1* locus (Sun et al., 2001).

Two of the best characterized examples of true allelic series are the downy mildew resistance gene *RPP13* in *Arabidopsis thaliana*, and the flax rust resistance gene *L*. Both show strong diversifying selection within the LRR region (Rose et al., 2004; Hall et al., 2009; Ravensdale et al., 2012), a domain involved in specificity and direct AVR protein binding (Dodds et al., 2006; Krasileva et al., 2010). Their allelic variants recognize naturally occurring variants of their cognate AVR, ATR13 and AVRL567, respectively, where differences in recognition specificity between the *R* alleles are associated with amino acid polymorphisms between the *Avr*-allelic variants (Sohn et al., 2007; Ellis et al., 2007). Both ATR13 and AVRL567 are typical candidate secreted effectors with an N-terminal secretory sequence. These effectors show no similarity to any characterized proteins, have no known biochemical function and no functional protein motif besides the RXLR translocation sequence in ATR13. The RXLR sequence was later found in many oomycete effectors (Wang et al., 2007; Sohn et al., 2007; Vleeshouwers et al., 2011). Different ATR13 and AVRL567 avirulence specificities are encoded within a single genetic locus or clustered, as are other AVR proteins of plant pathogens such as *Leptosphaeria maculans* (AVRLM1-2-6 and AVRLM3-4-7) and *Phytophthora infestans* (AVR3-AVR10-AVR11) (Allen et al., 2004; Dodds et al., 2004; Balesdent et al., 2002; van der Lee et al., 2001). However, not all *Avr* loci for an allelic series of resistance genes are necessarily clustered: for example, there are several genetically segregating *Avr* loci involved in the recognition by different alleles of the flax rust *L* resistance (reviewed by Ellis et al., 2007). In addition, and contrasting with the classical *Avr-R* interaction model, avirulence towards a few alleles of the *Mla* series was reported to be controlled by two independently segregating *Avr* loci in the pathogen. Avirulence for *Mla7* and *Mla13* was found to be controlled by two genes with unequal contributions, whereas two genes individually sufficient for avirulence on *Mla6* have been described (Brown and Simpson, 1994; Brown and Jessop, 1995; Brown et al., 1996; Caffier et al., 1996). Also, in flax rust, genetic evidence of an unlinked inhibitor of avirulence gene (*I*), inhibiting the recognition of several flax *Avr* genes (*AvrL567, AvrM1, AvrL1, AvrL8,* and *AvrL10*) by their corresponding resistance genes, was reported. However, the identity of this inhibitor gene and the mechanism that inhibits the recognition of these *Avrs* by their cognate *R* gene, are unknown (Lawrence et al., 1981; Jones, 1988a; 1988b; Ellis et al., 2007). Further evidence for genetic interactions which go beyond the simple gene-for-gene

model was found in the interaction of *Fusarium oxysporum* with the *I* resistance genes in tomato. The *Avr1* effector induces resistance in the presence of the *I-1* gene, but suppresses the resistance function of the *I-2* and *I-3* resistance genes (Houterman et al., 2008). In the same pathosystem, two additional specific suppressors of *I* gene function were identified in the pathogen, indicating the requirement of two fungal genes for immunity conferred by one resistance gene (Gawehns et al., 2014; Ma et al. 2015).

Powdery mildews are agronomically important and widespread pathogens of crop plants. In cereals, the two most studied sources of resistance against powdery mildew are the wheat *Pm3* and barley *Mla* allelic series, which encode CC-NBS-LRR type proteins (Yahiaoui et al., 2004; Seeholzer et al., 2010). The17 functional alleles of *Pm3* identified in cultivated hexaploid wheat are true alleles and share particularly high sequence identity (>97%) (Bhullar et al., 2009; 2010). In this study, we focus on the six alleles *Pm3a* to *Pm3f*. They have all been identified in classical breeding studies to confer different resistance specificity and have been shown to be allelic by molecular cloning (Yahiaoui et al., 2004; 2006; Srichumpa et al., 2005). In comparison, the *Mla* alleles share 84.6% sequence identity and show a similar divergence as the *A. thaliana RPP13* and the flax *L* alleles (Seeholzer et al., 2010; Rose et al., 2004; Ravensdale et al., 2012). Interestingly, some *Pm3* alleles were shown to confer overlapping but distinct resistance specificities. Specifically, the spectrum of powdery mildew races recognized by the *Pm3f* and *Pm3c* alleles is completely included within the larger spectrum of races recognized by *Pm3a* and *Pm3b,* respectively. They constitute pairs of 'weaker' (*Pm3f* and *Pm3c*) and 'stronger' (*Pm3a* and *Pm3b*) alleles sharing higher sequence identity among the *Pm3* alleles (Brunner et al., 2010). This is similar to the overlapping specificities of the *L5* and *L6* alleles in flax. However, although *L6* recognizes more alleles of *AvrL567*, the subset of alleles recognized by *L5* is not encompassed within the subset recognized by *L6* (Ravensdale et al., 2012). Stirnweis et al. (2014a) subsequently found that two amino acid changes within the Pm3 ARC2 subdomain of the NBS are sufficient to extend the resistance spectrum of the 'weaker' *Pm3f* allele to that of the 'stronger' *Pm3a* allele, providing evidence for the concept that specificity of R proteins is the sum of avirulence protein recognition plus activation of the resistance protein (Steinbrenner et al., 2015). This translates into a quantitative variation in the level of virulence in natural isolates, as described by Brunner et al. (2010) when they calculated haustorium index of natural isolates during transient expression assays of several *Pm3* alleles. Further investigation of inter-allele interactions revealed that the resistance mediated by the *Pm3a/f* alleles is suppressed by the *Pm3b/c* alleles, which suggests a complex mechanism underlying allelic race-specificity (Stirnweis et al. 2014b). Functional studies of the LRR domain indicated that very few amino acid polymorphisms are responsible for differences in allelic specificity (Brunner et al., 2010), thus suggesting direct recognition of allelic variants of an AVR effector, similarly to ATR13 and AVRL567.

Powdery mildew genomes contain a very large set of typical putative effectors based on their encoding small secreted proteins (Spanu et al., 2010; Wicker et al., 2013), which represent characteristic candidate AVR factors in fungi and oomycetes. Despite the multitude of resistance genes effective against *Blumeria graminis*, only two avirulence factors have been cloned to date: AVR_{a10} and AVR_{k1} from barley powdery mildew (Ridout et al., 2006). These *Avrs* are associated to non-LTR retrotransposons, and the mechanism of their interaction with *Mla10* and *Mlk1* is still unknown, as is the molecular basis of allelic specificity of the *AvrMla-Mla* interaction (Ridout et al., 2006).

Given the unique evolutionary history of the recently diverged allelic series of the *Pm3* resistance gene (Yahiaoui et al., 2009), combined with the evidence of complex inter-allelic interactions determining race-specificity (Brunner et al., 2010; Stirnweis et al., 2014a; 2014b), the *AvrPm3-Pm3* interaction provides a unique system to study the genetic and molecular mechanisms controlling specificity on the host as well as the pathogen side. In addition, evidence of recent evolution of wheat powdery mildew towards host specificity that is associated with a strong expansion of the effector repertoire (Wicker et al., 2013), provides a unique system to study host-pathogen co-evolution in light of the *Avr-R* interaction. In this context, identification of the *Avr* partners of the *Pm3* alleles is a prerequisite.

Here we show that *Pm3* multi-allelic race-specific resistance is controlled by multiple and genetically interacting loci in the fungus, where the final outcome depends on the genetic combination of avirulence genes and avirulence gene suppressors, a very different system from the expected simple *Avr* allelic series encoded within a single locus in the pathogen. We identify the *AvrPm3a/f* effector gene in wheat powdery mildew and demonstrate that it confers specificity of recognition in an allele-dependent manner. This demonstrates that powdery mildew putative secreted effectors can be AVR factors. Finally, we relate to previous studies of host *Pm3* allelic specificity and demonstrate that distinct resistance specificities are genetically controlled by distinct *Avr* loci in the fungus. We also demonstrate that suppressing and suppressed *Pm3* alleles recognize distinct *Avr* factors. Our findings show that the wheat powdery mildew pathosystem is highly complex both genetically and molecularly, adding novel insight into evolutionary aspects of *Avr-R* interaction and host-pathogen adaptation.

3.2 Results

Avirulence towards Pm3 powdery mildew resistance alleles is controlled by multiple pathogen loci

Very few genetic studies of the inheritance of avirulence/virulence in obligate biotrophic fungi have been carried out. The most extensive ones are those of flax rust (Flor, 1956; Lawrence et al., 1981) and barley powdery mildew (Brown and Simpson, 1994; Brown and Jessop, 1995; Brown *et al*., 1996; Caffier *et al*., 1996), which reported several examples where avirulence/virulence was genetically controlled by two loci in the fungus. We analyzed 167 F_1 haploid progeny from a cross between the *B.g. tritici* isolate 96224 that is avirulent on six different alleles of the *Pm3* resistance gene (*Pm3a, Pm3b, Pm3c, Pm3d, Pm3e* and *Pm3f*) and the isolate 94202 that is virulent on all these alleles. These progeny were phenotyped completely or partially on wheat differential lines carrying one of six different alleles of the *Pm3* resistance gene (*Pm3a*, *Pm3b*, *Pm3c*, *Pm3d*, *Pm3e* and *Pm3f*). They segregated for three classes of avirulence/virulence phenotypes. Based on the percentage of leaf coverage (LC) by the pathogen, progeny were scored either as avirulent (A, LC=0), intermediate avirulent (I, LC=10-40%) or virulent (V, LC=60-100%) (see Methods, see Appendix I: Supplemental Figure 1). Parlange et al. (2015) described a cross between the same avirulent parent 96224 and isolate JIW2 that is virulent on *Pm3c* and *Pm3f*. There, the F_1 population segregated for avirulence on *Pm3c* and *Pm3f* and three phenotypical classes were observed for *Pm3c*, whereas only two classes were observed for *Pm3f*. Data from this analysis will be referred to when relevant.

In the haploid F_1 progeny, each genetic locus from the parental genotypes is expected to segregate in a 1:1 ratio, so that only one of the two parental alleles is present per locus and per individual. Therefore, when translating phenotype segregation into "*Avr*" and "*avr*" genetic loci in the haploid mildew, it is important to consider the following hypotheses as equally valid: i) avirulence/virulence polymorphism is explained by the *Avr* allele inherited from the avirulent parent, which encodes a factor that promotes avirulence, or ii) avirulence/virulence is explained by the *avr* allele inherited from the virulent parent, which encodes a factor that promotes virulence. We used the precedents of *B.g. hordei* where intermediate phenotypes and complex segregation ratios were reported (Brown and Simpson, 1994; Brown and Jessop, 1995; Brown *et al*., 1996; Caffier *et al*., 1996), to formulate and test genetic models for the control of avirulence/virulence in *B.g. tritici*.

Results of the genetic analysis are summarized in Figure 1, and detailed procedures for phenotype segregation analysis as well as phenotype-genotype assignments are given in Methods, and in Appendix I: Supplemental Figure 2 and Supplemental Text. On wheat genotypes containing *Pm3a*, *Pm3c*, and *Pm3e*, phenotype segregation fitted a 1:1

ratio (A:V) (Table 1), consistent with a model where the interaction is controlled by a single genetic locus differing between avirulent and virulent parents (Figure 1A). A progeny receiving the allele from the avirulent parent at locus *A1/a1*, *C1/c1* and *E1/e1* is avirulent on *Pm3a*, *Pm3c* and *Pm3e*, respectively. By contrast, a progeny receiving the allele from the virulent parent at locus *A1/a1*, *C1/c1* and *E1/e1* is virulent on *Pm3a*, *Pm3c* and *Pm3e*, respectively. Whether this locus controls virulence or avirulence cannot be determined at this stage of the analysis but will be considered later.

Segregation on *Pm3b*, *Pm3d* and *Pm3f* did not fit the one locus segregation model, suggesting there is more than one locus controlling the avirulent/virulent phenotype and distinguishing the avirulent and virulent parents with respect to these *Pm3* alleles. On *Pm3f*, phenotype segregation fitted a ratio of 1 avirulent to 3 virulent, that is consistent with a model where two independently segregating loci differentiating the parents are involved in the interaction (*F1/f1*, *F2/f2*). This model predicts that avirulence/virulence is controlled by an interaction between these two loci, where only progeny with the genotype of the avirulent parent at both loci are avirulent on *Pm3f* (Figure 1B). Segregation on *Pm3b* and *Pm3d* included the third class of intermediate phenotypes, and did not fit the single locus or the two loci segregation models. On *Pm3b*, the F_1 segregation fitted a 5:1:2 ratio (A:I:V; Table 1), consistent with a model where three independently segregating loci differentiating the parents are involved in the interaction (*B1/b1, B2/b2, B3/b3*). Based on the ratio of 5:1:2 (A:I:V), this model predicts that progeny receiving the avirulence allele at locus *B1/b1* are avirulent, regardless of the genotype at the other two loci. Progeny with the virulence allele at locus *B1/b1* which have the avirulence alleles at both *B2/b2 and B3/b3* loci are also avirulent. Progeny with only the avirulence allele from locus *B2/b2* show intermediate avirulence. Individuals that possess only the avirulence allele at locus *B3/b3* and individuals with no avirulence alleles are virulent (Figure 1C and Appendix I: Supplemental text). Finally, segregation on *Pm3d* fitted a 2:1:5 (A:I:V) ratio (Table 1), consistent with a model involving three independently segregating loci differentiating the parents (*D1/d1, D2/d2, D3/d3*). Based on the ratio of 2:1:5 (A:I:V) this model predicts that progeny receiving the avirulence alleles at both loci *D1/d1* and *D2/d2* are avirulent. Progeny receiving the avirulence alleles at loci *D2/d2* and *D3/d3* show an intermediate phenotype. All other progeny are virulent (Figure 1C and Appendix I: Supplemental text).

Figure 1. Interpretation of genetic segregation in the F1 progeny of the cross between the wheat powdery mildew isolates 96224 and 94202. The genetic loci segregating for avirulence in the population are color coded. *Locus_1*, involved in the six *AvrPm3-Pm3* interactions, is indicated in "red". *Locus_2*, involved in the *AvrPm3^{f2}-Pm3f* interaction is indicated in "purple". *Locus_3*, involved in the *AvrPm3^{b2}-Pm3b*, *AvrPm3^{c2}-Pm3c* and *AvrPm3^{d2}-Pm3d* interactions, is indicated in "blue". Finally, the unmapped *AvrPm3^{b3}* and *AvrPm3^{d3}* loci are indicated in "orange" and "green", respectively. At each locus, the genotypes from the avirulent 96224 and virulent 94202 parents are indicated by darker and lighter shades, respectively. For simplicity the *AvrPm3* loci were abbreviated by a single letter indicating the genotype of the avirulent (uppercase) and virulent (lowercase) parent. Phenotypes are indicated by A (avirulence), I (intermediate avirulence) and V (virulence). Segregation ratios of A:I:V are indicated. Detailed description and interpretation of phenotype segregation ratios from subfigures (A), (B) and (C) are provided in the Results section and Appendix I: Supplemental Text. (a) Avirulence towards *Pm3a*, *Pm3c* and *Pm3e* is determined by the parental genotype inherited at a single locus, which is commonly involved in the six *AvrPm3-Pm3* interactions. (b) Avirulence towards *Pm3f* is determined by the parental genotype inherited at two loci. One out of the four possible genotype combinations determines avirulence on *Pm3f*, and three determine virulence. (c) Avirulence towards *Pm3b* and *Pm3d* is determined by the parental genotype inherited at three loci. Eight genotype combinations are possible. For *Pm3b*, five determine avirulence, one determines intermediate avirulence and two determine virulence. For *Pm3d*, two combinations determine avirulence, one determines intermediate avirulence, and five determine virulence.

Table 1: Segregation of race-specific avirulence in the F_1 mapping population

Pm3 resistance		Progeny phenotype[b]			Genetic segregation		
Alleles	Cultivar/line[a]	A	I	V	Ratio[c]	Chi²	*P*[d]
Pm3a	Asosan^{8*CC}	81	0	85	1:0:1	0.096	0.756
Pm3b	Chul^{8*CC}	108	17	33	5:1:2	2.319	0.313
Pm3c	Sonora^{8*CC}	77	1	81	1:0:1	0.107	0.743
Pm3d	Kolibri	44	18	96	2:1:5	0.744	0.689
Pm3e	W150	35	1	32	1:0:1	0.147	0.701
Pm3f	Michigan Amber^{8*CC}	46	2	119	1:0:3	0.745	0.388

[a] Wheat cultivars/lines carrying the *Pm3* allele indicated in the first column. Near isogenic lines obtained after backcrossing eight times in cultivar "Chancellor", are indicated by the superscript "8*CC".
[b] Progeny phenotype was scored as leaf coverage (LC). Phenotypes are indicated "A" for avirulent (LC=0), "I" for intermediate avirulent (LC=10-40%) and "V" for virulent (LC=60-100%).
[c] Genetic segregation is given as a theoretical ratio of A:I:V. Deviation of phenotype numbers from the theoretical ratios was tested with the Chi² test for goodness of fit. The degree of freedom was assigned as the "*Number of phenotypic classes – 1*".
[d] The probability value for the Chi² test. $P<0.05$ would indicate significant deviation from the theoretical ratios.

We have also compared phenotype patterns in the 143 F1 progeny for which we had complete phenotype data on five alleles of *Pm3* (Figure 2A). As previously mentioned, a significant proportion of the progeny segregated for intermediate avirulence on *Pm3b* and *Pm3d*. However, the subsets segregating for this phenotype were different between these two alleles (Figure 2A), suggesting that different loci or loci combinations in the pathogen are responsible for conferring the intermediate phenotype towards *Pm3b* and *Pm3d*. We found that $^1/_8$ of the progeny was consistently avirulent on *Pm3a/b/c/d/f* and probably also on *Pm3e* although we had only partial phenotypic data on this allele (Figure 2A). We also found that the subset of progeny that is avirulent on *Pm3b* included all the progeny found to be avirulent on the six *Pm3* alleles (Figure 2A). This pattern is best explained by the presence of a genetic factor in the fungus commonly involved in the interaction with the six alleles of *Pm3*. Overall, these data indicate a complex genetic control of avirulence that could only be addressed by mapping the loci controlling the *AvrPm3-Pm3* interaction across the genome.

Genome-wide mapping of the genetic loci controlling the *AvrPm3-Pm3* interaction

In addition to the F_1 progeny from the cross between 96224 and 94202 (hereafter referred to as the 94202 population), we have previously produced F_1 progeny from a cross between 96224 and the British isolate JIW2 (hereafter referred to as the JIW2 population).

The JIW2 population segregated only for *AvrPm3f* and *AvrPm3c* (Parlange et al., 2015), but unlike the 94202 population, the F_1 segregation indicated that only one locus genetically controls avirulence towards *Pm3f*, while two loci are genetically involved in conferring avirulence and intermediate avirulence towards *Pm3c* (only one locus segregates in the 94202 population for *Pm3c*). To map the loci controlling avirulence towards *Pm3* in both populations, we chose to use progeny from both crosses for genome-wide SNP mapping. The genome sequence of the three parental isolates is available (Wicker et al., 2013), and SNPs were identified by mapping Illumina sequence reads from 94202 and JIW2 on to the 96224 reference genome. A total of 164 F_1 progeny from the JIW2 population, and 154 from the 94202 population, were scored for the presence of 254 and 228 SNP loci, respectively, using Kompetitive Allele-Specific PCR (KASP) technology (He et al., 2014). By design, 200 of these SNPs corresponded to identical polymorphisms between 94202 and JIW2 as compared to the reference 96224, allowing comparative mapping between populations. Of all the loci scored with KASP, 251 (98.8%) in the JIW2 population and 224 (98.2%) in the 94202 population were found to be polymorphic and were used to establish a genetic map.

We generated two genetic maps, using SNP-derived KASP markers, and an additional 80 AFLP markers were available for the JIW2 population (Parlange et al., 2015). Both maps consisted of 17 linkage groups (LGs) of at least 4 markers, with a total size of 1432cM for the JIW2 map and 1227cM for the 94202 map (see Appendix I: Supplemental Figures 3 and 4). Because of the markers shared between the individual maps, it was possible to produce three consensus LGs containing all the loci controlling the *AvrPm3-Pm3* interactions mapped in both populations (Figure 2B, see Appendix I: Supplemental Figure 4). Consistent with the observation that a subset of progeny are avirulent on all *Pm3* alleles, we mapped one locus in the fungus genetically interacting with the six tested alleles (*locus_1*) (see Appendix I: Supplemental Figure 5 and Supplemental Text). By selecting different subsets of progeny, we mapped a second locus genetically interacting with *Pm3f* (*locus_2*), and a third locus genetically interacting with *Pm3b*, *Pm3d* and *Pm3c* (*locus_3*) (see Appendix I: Supplemental Figure 5 and Supplemental Text). Due to insufficient number of progeny, we were not able to map the genetic regions corresponding to the loci *B3/b3* and *D3/d3*, which genetically interact with *Pm3b* and *Pm3d*, respectively (Figure 1C).

These data suggest that *locus_1* might encode a general factor involved in all *AvrPm3-Pm3* interactions. We hypothesized that specificity towards different alleles of *Pm3* is likely to be controlled by the genotype of the avirulent parent at *locus_2* for *Pm3f* (*AvrPm3^{f2}*) and *locus_3* for *Pm3b, Pm3c, Pm3d* (*AvrPm3^{b2}, AvrPm3^{c2}, AvrPm3^{d2}*). For *locus_2,* which specifically interacts with the *Pm3f* allele, we hypothesized that only one genetic factor determining allelic-specificity is likely to be encoded within this locus in the avirulent parent.

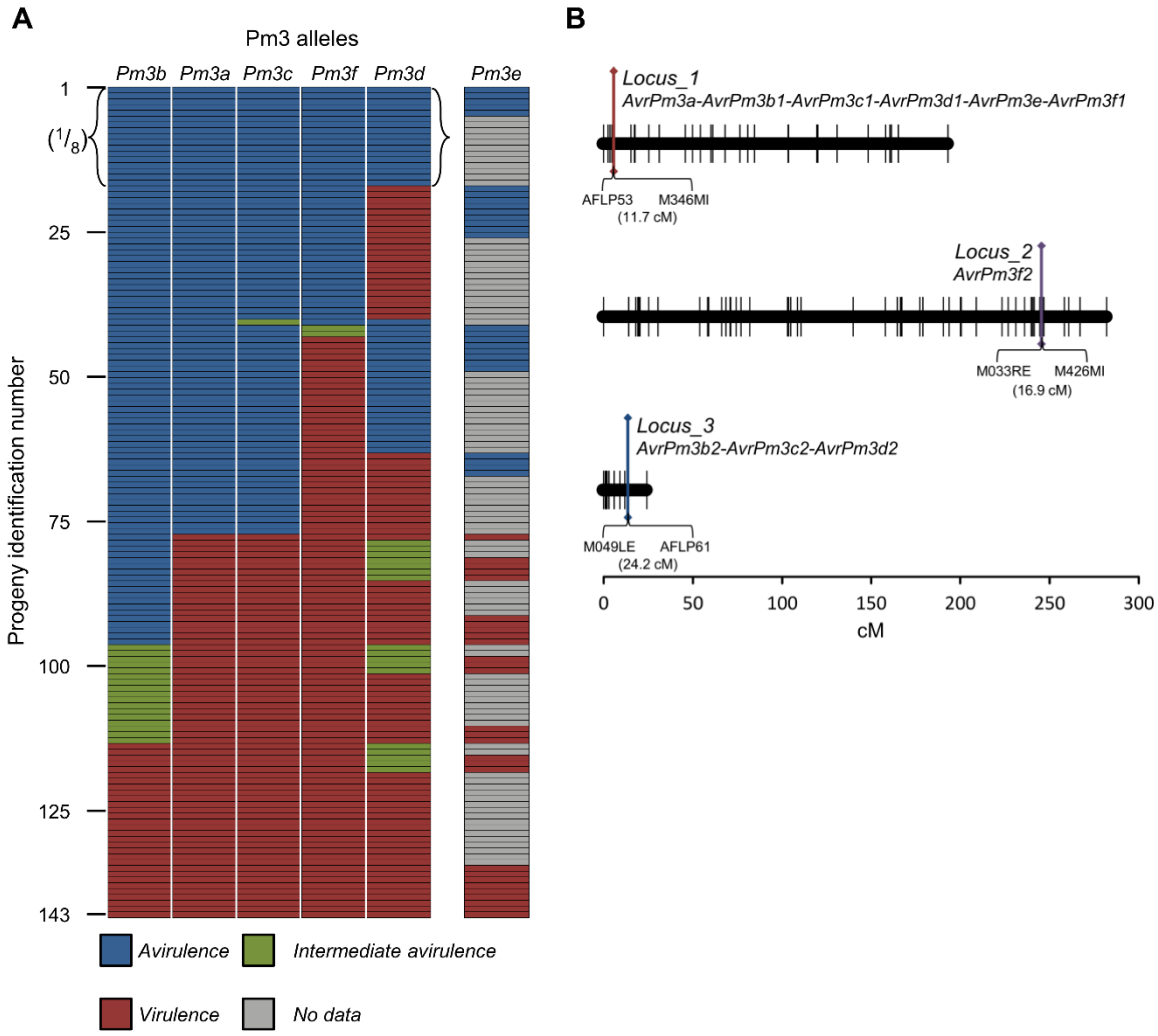


Figure 2. Identification and mapping of three major loci involved in the *AvrPm3-Pm3* interaction. (A) Comparative analysis of the reaction of individual progeny isolates from an F_1 segregating population of the cross 96224 x 94202 on the different *Pm3* resistance alleles. Phenotypes on six *Pm3* alleles are indicated for 143 progeny and sorted by avirulence (blue), intermediate avirulence (green) and virulence (red). The identity of the progeny is indicated according to phenotype sorting, by an identification number on the left side. A subset of 17 progeny (indicated with braces) corresponding to about $^1/_8$ of the population is consistently avirulent on five *Pm3* alleles. The subset of avirulent progeny on *Pm3b* includes all avirulent progeny on the five other alleles. The subset of progeny showing intermediate avirulence on *Pm3b* is different from the one showing this phenotype on *Pm3d*. (B) Linkage groups of consensus maps of the wheat powdery mildew genome containing loci involved in the *AvrPm3-Pm3* interaction. The genotype of the avirulent parent at these loci is indicated. Three loci were mapped on three consensus linkage groups. The consensus map was produced by merging the linkage groups containing the *AvrPm3/avrPm3* specificities, individually mapped in different progeny subsets, or across two populations (see Appendix I: Supplemental Figures 3, 4, 5, and 6, and Supplemental Text). Co-localization of different *AvrPm3/avrPm3* loci was assigned based on the closest genetically linked KASP or AFLP marker and comparison of phenotype patterns in the progeny. Flanking markers inferred from consensus mapping are indicated. Genetic distance between the flanking markers as well as the size of the linkage group, are indicated in centiMorgan (cM).

By contrast, for *locus_3*, which genetically interacts with three different *Pm3* alleles (*Pm3b*, *Pm3c* and *Pm3d*), we hypothesized that up to three different genetic factors determining allelic-specificity are likely to be encoded in this locus in the avirulent parent. Thus, because our data suggest a complex genetic control of avirulence by *locus_3* (see Appendix I: Supplemental Text), and because we wanted to identify an allele-specific avirulence gene, we chose to focus our efforts on high-resolution mapping and subsequent map-based cloning of *AvrPm3^{f2}* in *locus_2*.

High-resolution mapping of *AvrPm3^{f2}* in locus_2

To map the genetic interval containing *AvrPm3^{f2}*, we used a subset of 70 F_1 progeny from the 94202 population that are all avirulent on *Pm3a* and *Pm3c*, thus indicating they all have the genotype of the avirulent parent 96224 at *locus_1* (*AvrPm3^{f1}*) according to the 1:1 (A:V) phenotypic segregation ratio on *Pm3a* and *Pm3c.* This subset of progeny segregates in a 1:1 ratio (A:V) on *Pm3f*, allowing to distinguish *AvrPm3^{f2}* and *avrPm3^{f2}* (see Appendix I: Supplemental Figure 6). Using KASP genotyping data from the selected 70 progeny, we generated a low resolution genetic map on which we mapped *AvrPm3^{f2}* within a genetic interval of 20cM between the markers M033RE and M426MI (Figure 3A). To further reduce the genetic interval, we used whole genome sequencing of F_1 progeny in a process similar to Bulk Segregant Analysis (BSA) (Takagi et al., 2013, see Appendix I: Supplemental Figures 6 and 7), and a series of Cleaved Amplified Polymorphic Sequence (CAPS) markers designed on SNPs between the parental isolates. We produced the left flanking marker 33M11, mapping at 0.7 cM (1 recombinant) from *AvrPm3^{f2}* on FPC contig-33, and the right flanking marker 52M2 mapping at 1.5 cM (2 recombinants) from the *Avr* on FPC contig-52 (Figure 3B, 3C). In the powdery mildew genome assembly, no overlap was predicted between the flanking FPC contigs 33 and 52 in the physical map.

Sequencing, assembly and annotation of locus_2

To uncover the full sequence of the locus, we chose to sequence the minimal tiling path of the Bacterial Artificial Chromosome (BAC) contigs carrying both flanking markers of *AvrPm3^{f2}*. BAC fingerprints initially assembled using the FPC algorithm (Figure 3B) (Wicker et al., 2013), were assembled *de novo* using the Linear Topological Contig (LTC) algorithm (Frenkel *et al.*, 2010), which is specifically designed to assemble BACs from highly repetitive genomes. In the LTC assembly, clusters that contained the BACs of contig-33 and contig-52 could be linked together to form scaffold 10 (see Appendix I: Supplemental Figure 8). Physical markers designed to cover the predicted BAC overlaps confirmed the tiling path, and the 7 most informative BAC clones were individually Illumina sequenced (Figure 3D). Sequencing reads were assembled into a 445 kb contiguous sequence that contained mainly transposable elements (90% of the sequence)

and fourteen genes (Figure 3E). Among these, eight were predicted to encode putative secreted effector proteins. The remaining six annotated genes had no predicted protein function.

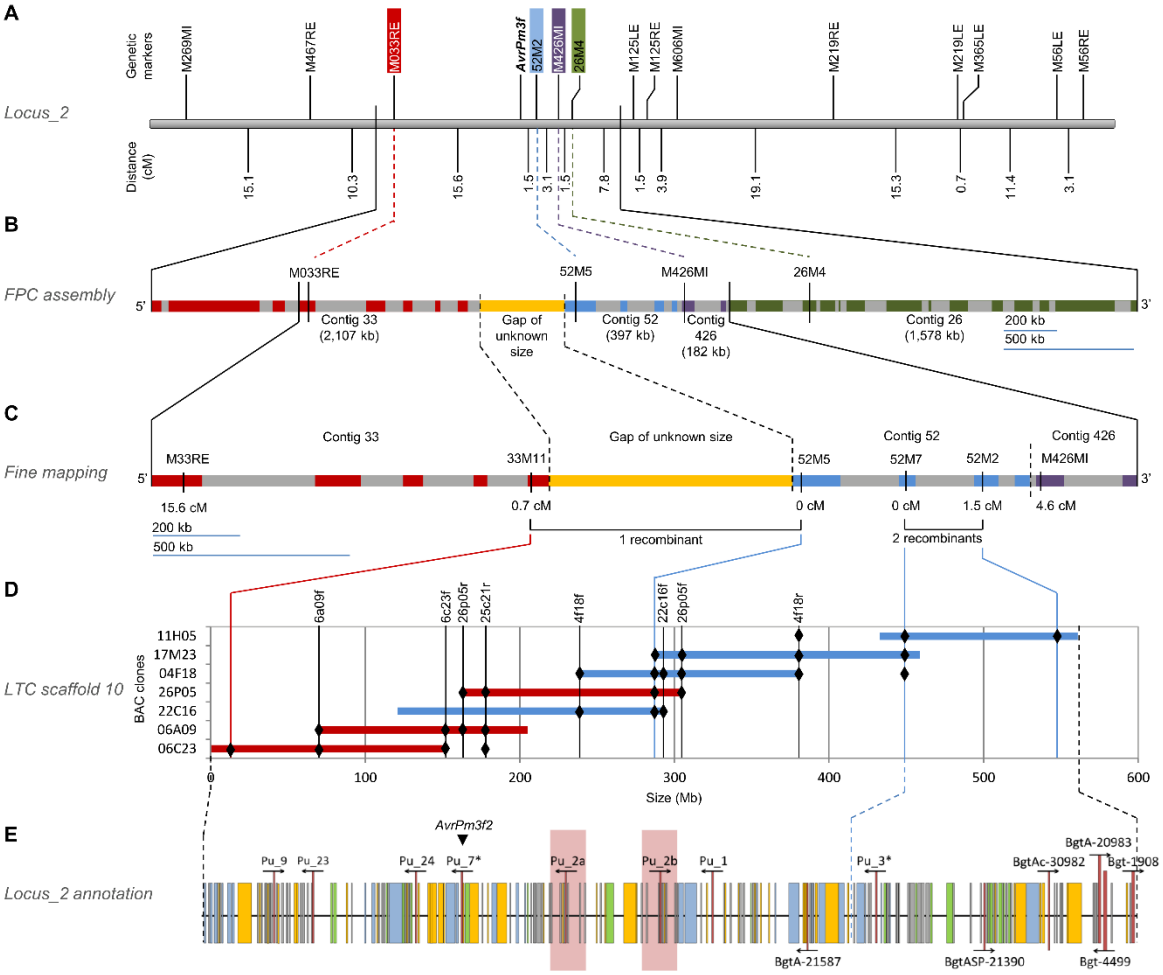


Figure 3. High-resolution mapping of *AvrPm3*^{f2} in *locus*_2. For each level, the most informative genetic or physical interval was chosen to be depicted and the 5′ to 3′ orientation is arbitrarily assigned to improve clarity, in the absence of knowledge of the location of the centromere. (A) Linkage group 2 of the genetic map containing the *AvrPm3*^{f2} locus, the two closest KASP markers, M33RE (red) and M426MI (purple), the flanking marker on contig-52, 52M2 (blue), and a marker for contig-26, 26M4 (green). Distances are indicated in centiMorgan (cM). (B) Fingerprinted Contigs (FPCs) covering the *AvrPm3f2* region. Contigs linked to *AvrPm3*^{f2} by the genome-wide SNP genotyping approach (contig-33 and -426) and by bulk segregant analysis (contig-52 and -26) are shown, their estimated sizes and location of genetically informative markers are given. Red, blue, purple and green boxes denote known sequence, while grey boxes indicate sequence gaps as derived from the wheat powdery mildew genome sequence (Wicker et al., 2013). Yellow is a sequence gap of unknown size in the FPC assembly. (C) Fine genetic mapping in the *AvrPm3*^{f2} region. The most informative CAPS markers and the two flanking KASP markers are shown, and their genetic distances from *AvrPm3*^{f2} are provided. The number of recombinants between the flanking and co-segregating markers is specified. (D) BAC minimum tiling path of LTC scaffold 10, spanning the *AvrPm3*^{f2} region. Markers are indicated by vertical bars and specific PCR amplifications from both CAPS and BAC-end markers are indicated by black diamonds. Horizontal bars indicate BACs from FPC contig-33 (red) and FPC contig-52 (blue). Only the informative BACs selected for sequencing are shown. (E) Annotated *locus_2* sequence. Putative effector genes and their transcriptional orientation are given. Polymorphic genes within the candidate interval are indicated (asterisk). The functionally validated *AvrPm3*^{f2} is indicated by a black arrow. An inverted duplication of 15 kb is highlighted (red overlay). Transposable elements are depicted as colored boxes.

Identification and functional validation of *AvrPm3$^{a2/f2}$*

To identify the avirulence gene candidates in *locus_2*, we searched for genes that were polymorphic between the parental isolates and genetically co-segregated with *AvrPm3^{f2}*. Among the fourteen genes predicted on *locus_2* only *Pu_3* and *Pu_7* were polymorphic between 96224 and 94202 sequences, and the polymorphisms were validated by re-sequencing of specific PCR amplicons of the two genes from genomic DNA. *Pu_3* was absent in the virulent parent 94202, while *Pu_7* showed two SNPs between the parental sequences, one in each of the two predicted exons (Figure 4A). Closer analysis revealed that both SNPs lead to amino acid polymorphisms in the encoded protein. The first SNP results in a non-conservative amino acid substitution from the aliphatic, non-polar glycine in the avirulent parent to acidic, polar glutamate in the virulent one. The second SNP results in a conservative amino acid substitution from alanine in the avirulent parent to valine in the virulent one. To test for co-segregation of *Pu_7* with *AvrPm3^{f2}*, we relied on a SNP in the gene coding sequence distinguishing between the parental alleles, to derive a gene specific CAPS marker (pu7_CAPS, see Appendix I: Supplemental Table 1). This marker was tested on the subset of progeny segregating in a 1:1 ratio for *AvrPm3^{f2}*, and mapped at 0 cM from the *Avr*, which demonstrate that *Pu_7* co-segregates with *AvrPm3^{f2}*. For *Pu_3*, given its proximity to marker 52M7, which co-segregates with *AvrPm3^{f2}*, and its presence-absence polymorphism between the parents, we decided to consider it also as a valid *Avr* candidate.

Brunner et al. (2010) previously demonstrated that *Pm3f* and *Pm3a* are two alleles with very characteristic resistance spectra: based on the observation that all the isolates recognized by *Pm3f* were also recognized by *Pm3a*, but not vice versa, they concluded that *Pm3a* is a 'stronger' form of *Pm3f*. Later, Stirnweis et al. (2014a) showed that a substitution of only two amino acids in the NBS domain of PM3F (PM3F $^{L456P/Y458H}$) was sufficient to broaden *Pm3f*-mediated resistance to potentially confer the same race-specificity as *Pm3a*, thus substantiating the hypothesis that *Pm3a* is a 'stronger' version of *Pm3f*. Therefore, to test for interaction with the resistance gene, *Pu_3* and *Pu_7* were cloned from the avirulent parent 96224 (*Pu_3^{avr}, Pu_7^{avr}*) and the virulent parent 94202 (*Pu_7^{vir}*). Constructs expressing protein versions with and without signal peptide were transiently co-expressed by agroinfiltration in *Nicotiana benthamiana* with the *Pm3f*, *Pm3f* $^{L456P/Y458H}$, and *Pm3a* resistance genes. No reaction was observed in any infiltrations involving the two versions of *Pu_3^{avr}* (with and without signal peptide) in combination with the three *Pm3* genes, which indicates that *Pu_3^{avr}* is not the *Avr*. For *Pu_7*, a strong hypersensitive cell death response was observed in all infiltrations where *Pu_7^{avr}* version without signal peptide (Figure 4B) was co-expressed with *Pm3f* $^{L456P/Y458H}$ and *Pm3a* (Figure 5A).

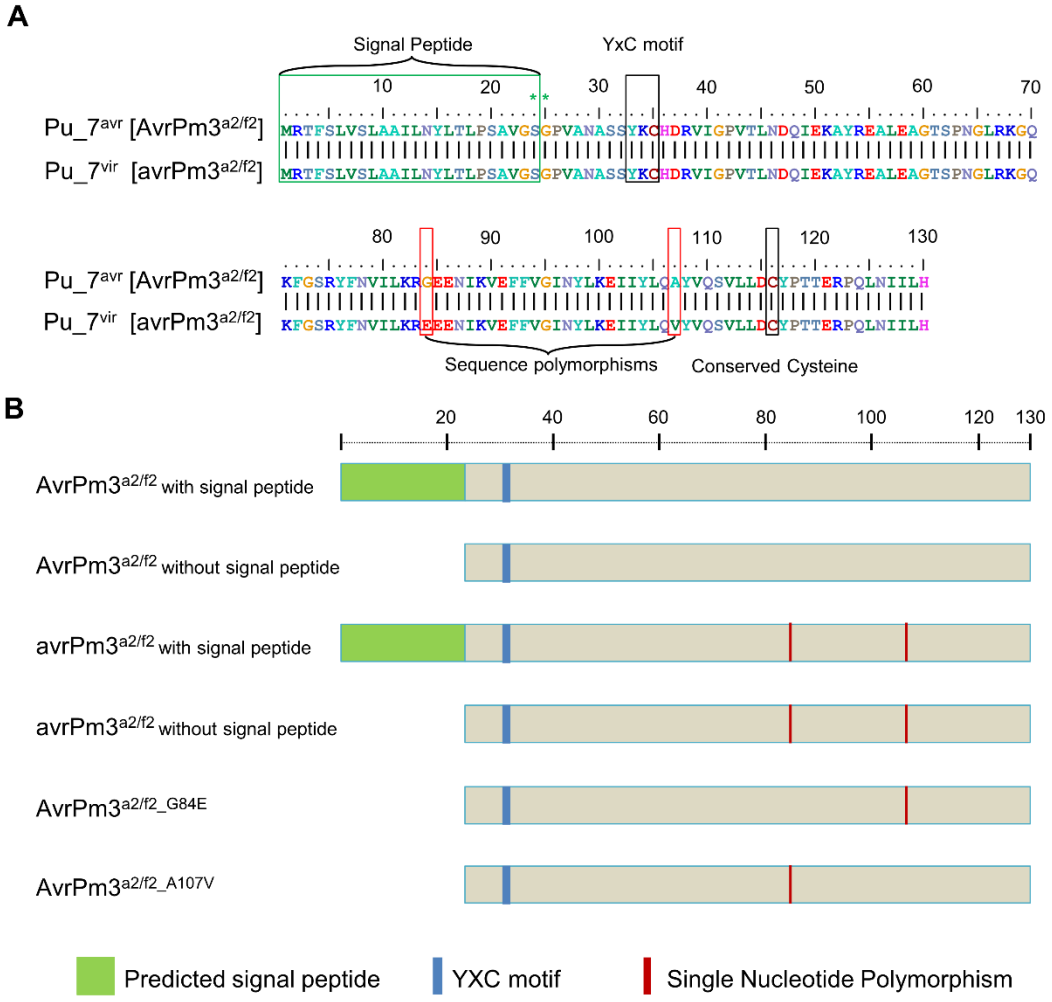


Figure 4. Description of the *AvrPm3$^{a2/f2}$* encoded proteins and protein variants used in this study. (A) Alignment of the two protein sequences encoded by the avirulence (*Pu_7^{avr}*) and the virulence alleles (*Pu_7^{vir}*) of the *Pu_7* effector gene. Amino acids are colored according to their biochemical properties. The predicted signal peptide (green box) and cleavage site (green asterisks) are indicated. Total protein size is 130 amino acids, with a predicted signal peptide of 24 residues, and a predicted mature protein of 106 residues. The YxC motif and the second conserved cysteine are indicated (black boxes). The two amino acid polymorphisms between the parental alleles, one non-conservative (Gly/Glu), one conservative (Ala/Val), are shown (red boxes). (B) Schematic diagram of the protein variants used in transient protein expression assays. At the top, the full length protein encoded by *AvrPm3$^{a2/f2}$* is shown (130 residues). The predicted signal peptide (SignalP v. 4.1) is indicated in green. The location of the YxC motif is indicated in blue. Red bars indicate single nucleotide polymorphisms distinguishing between virulence and avirulence alleles (*AvrPm3$^{a2/f2}$*, *avrPm3$^{a2/f2}$*), and the two variants generated by site directed mutagenesis (AVRPM3$^{A2/F2_G84E}$, AVRPM3$^{A2/F2_A107V}$).

We have also co-expressed this version of *Pu_7^{avr}* with the weak *Pm3f* allele, which resulted in weak HR in 23% of the leaves (see Appendix I: Supplemental Figure 9). These results indicate that *Pu_7^{avr}* confers dual specificity towards *Pm3f* and *Pm3a*, and further substantiate that the two amino acid polymorphisms within the nucleotide-binding

domain of PM3F are lowering the cell-death response upon AVR recognition. Therefore, for clarity and consistency with the genetics, *Pu_7^{avr}* is hereafter referred to as *AvrPm3$^{a2/f2}$*. No reaction was obtained in all infiltrations with the *Pu_7^{vir}* allele (*avrPm3$^{a2/f2}$*) (Figure 4B and 5A) demonstrating that *AvrPm3$^{a2/f2}$* behaves as a race-specific *Avr* with an avirulent version that is recognized by its cognate R protein, and a virulent version that escapes recognition.

In the infiltrations involving *AvrPm3$^{a2/f2}$*version with signal peptide (Figure 4B), in combination with *Pm3f*$^{L456P/Y458H}$ or *Pm3a*, a weak HR was observed as compared to the controls involving the version without signal peptide (Figure 6A). Our interpretation is that the removal of the signal peptide from AVRPM3$^{A2/F2}$ allows the protein to accumulate inside the cell instead of being secreted to the apoplast. Therefore, considering the putative intracellular localization of the PM3 proteins, these results suggest that AVRPM3$^{A2/F2}$ probably interacts with PM3A and PM3F $^{L456P/Y458H}$ in the cytosol or the nucleus.

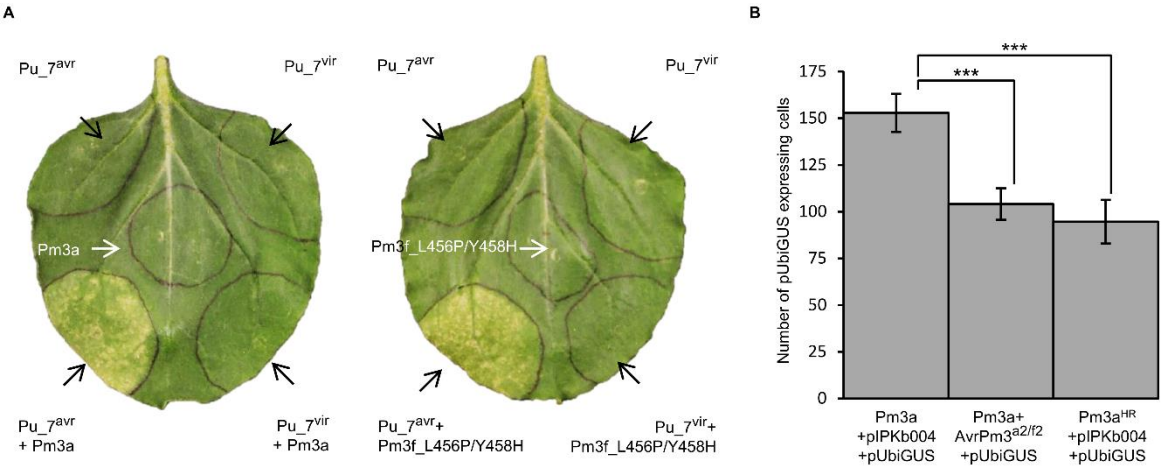


Figure 5. Functional validation of the *AvrPm3$^{a2/f2}$-Pm3a/f* interaction. (A) Agroinfiltration assays in *Nicotiana benthamiana* demonstrate induction of an HR response upon specific recognition of *Pu_7^{avr}* (*AvrPm3$^{a2/f2}$*) by *Pm3a* and *Pm3f$^{L456P/Y458H}$*. The avirulence (*Pu_7^{avr}*) and virulence (*Pu_7^{vir}*) alleles of *Pu_7* without signal peptide were transiently expressed together with *Pm3a* and *Pm3f$^{L456P/Y458H}$*. Leaves of 4 week old *N. benthamiana* plants were infiltrated with *Agrobacterium* cultures expressing each of the constructs indicated. An *Avr*:*R* ratio of 4:1 and OD_{600} of 1.2 were used to test for HR induction. Results are consistent across replicates from at least three independent experiments where 4 to 8 leaves were assayed. Photographs were taken 5 days after infiltration. (B) Co-bombardment of *AvrPm3$^{a2/f2}$* (*Pu_7^{avr}*) with *Pm3a* in wheat reduces the number of GUS expressing cells. The *AvrPm3$^{a2/f2}$* construct encoding for a protein version without signal peptide was transiently co-expressed in wheat leaf segments with *Pm3a* and the pUbiGUS reporter, using the particle bombardment assay. The number of GUS expressing cells was assessed 48 hours after bombardment and compared between leaves co-expressing: (i) *Pm3a* + *AvrPm3$^{a2/f2}$*, (ii) *Pm3a* + empty vector (pIPKb004) as a control and (iii) the cell death inducing auto-active *Pm3a^{HR}* + empty vector (pIPKb004) as a second control. Values are given as the average number of GUS stained cells per slide containing 5 leaves. Six slides were counted for each of 4 to 6 bombardment assays (see Methods). Statistical significance was assessed using the unpaired Student's t-Test. Significance under the threshold $P \leq 0.001$ (***) is indicated.

We then investigated whether *AvrPm3*$^{a2/f2}$was recognized by other alleles of *Pm3*. Using the same assay in *N. benthamiana*, *AvrPm3*$^{a2/f2}$was transiently co-expressed with *Pm3b*, *Pm3c*, *Pm3d* and *Pm3e*, and no HR was observed. This indicates that *AvrPm3*$^{a2/f2}$is specific to *Pm3a* and *Pm3f,* suggesting that there are other *AvrPm3* genes which specifically interact with the other *Pm3* alleles, which is in agreement with our observation described above that multiple loci in the fungus are involved in the *AvrPm3-Pm3* interaction.

To test for *AvrPm3*$^{a2/f2}$*-Pm3a/f* interaction in wheat, we used particle bombardment to transiently co-express *AvrPm3*$^{a2/f2}$and *Pm3a* in the susceptible wheat cultivar "*Chancellor*". We used the pUbiGUS reporter plasmid to assess β-glucuronidase (GUS) accumulation in leaf segments co-bombarded with *AvrPm3*$^{a2/f2}$and *Pm3a* as compared to (i) a negative control consisting of a co-bombardment of the empty expression vector (pIPKb004) with *Pm3a* and (ii) a positive control consisting of a co-bombardment of the empty expression vector with *Pm3a*^{HR}, an auto-activated version of *Pm3a* triggering constitutive cell death response (Stirnweis et al., 2014b). The number of GUS expressing cells was revealed by GUS staining, 48 hours after bombardment. The number of GUS expressing cells was significantly reduced when co-bombarding *AvrPm3*$^{a2/f2}$with *Pm3a*, as compared to the *Pm3a* negative control (Figure 5B). This reduction was as significant as with the cell-death inducing *Pm3a*^{HR} control, indicating that GUS accumulation was similarly compromised in cells co-expressing *AvrPm3*$^{a2/f2}$and *Pm3a*. These results in wheat corroborate the data obtained by transient expression assays in *N. benthamiana*, which demonstrate that *AvrPm3*$^{a2/f2}$is the effector gene partner of the wheat *Pm3* resistance gene alleles *Pm3a* and *Pm3f*.

To summarize, *AvrPm3*$^{a2/f2}$ co-segregates with *AvrPm3f* in *locus_2*. It is recognized by *Pm3f* (weaker HR is obtained with the natural allele in *N. benthamiana* as compared to the *Pm3f* $^{L456P/Y458H}$ variant), and by *Pm3a* (in transient assays in *N. benthamiana* and wheat), which demonstrates that *AvrPm3*$^{a2/f2}$ is *AvrPm3a* and *AvrPm3f*.

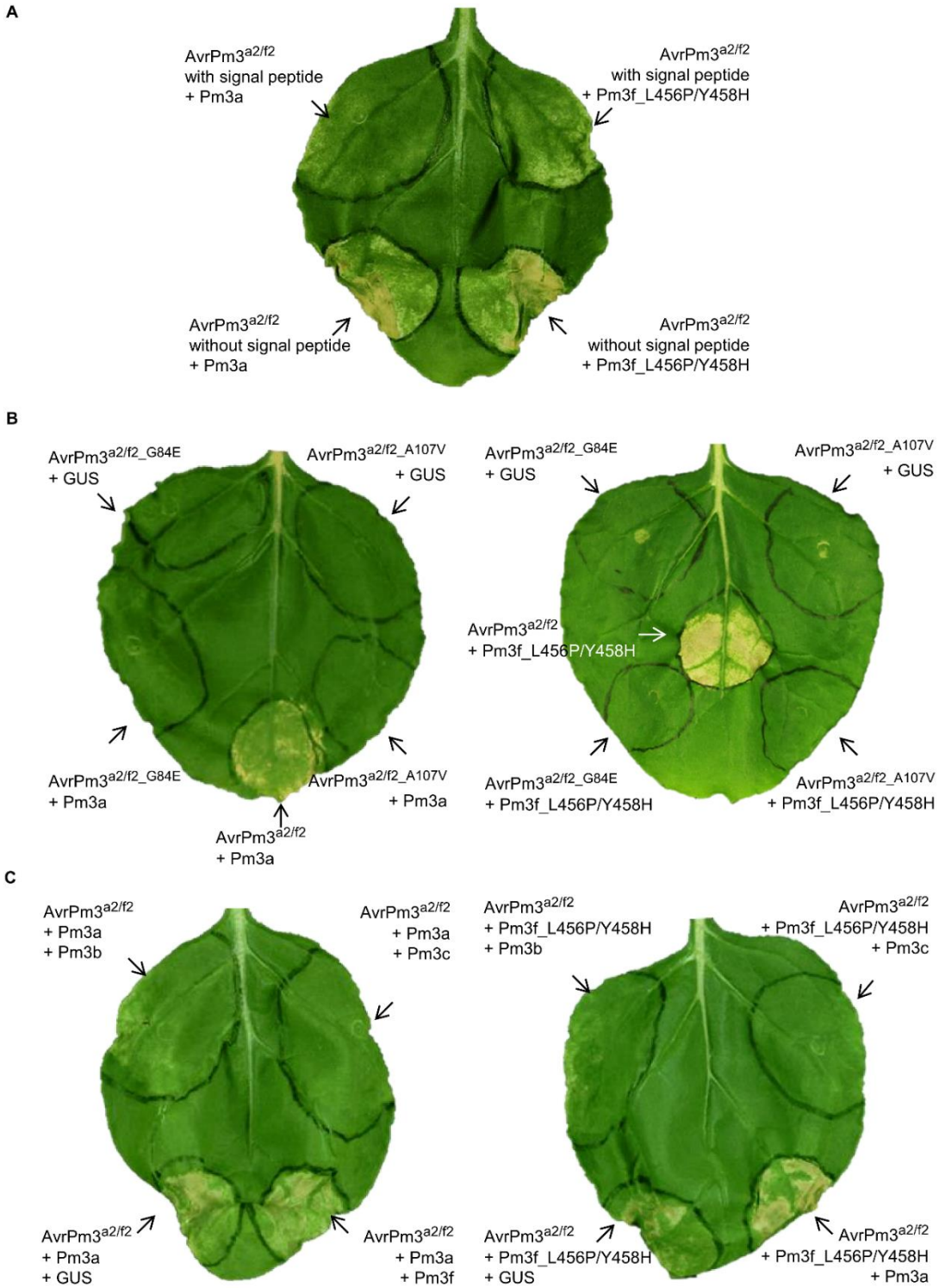


Figure 6. Functional analysis of the *AvrPm3$^{a2/f2}$-Pm3a/f* interaction in the *N. benthamiana* system.

(Figure 6. Continued from previous page) (A) Agroinfiltration assays in *N. benthamiana* with *AvrPm3$^{a2/f2}$* constructs encoding for the protein versions with and without signal peptide, co-infiltrated with *Pm3a* or *Pm3f$^{L456P/Y458H}$*. Stronger HR was observed with the *AvrPm3$^{a2/f2}$*construct expressing the version without signal peptide. Photographs were taken 5 days after infiltration. (B) Both polymorphic residues in the protein encoded by *AvrPm3$^{a2/f2}$*are necessary for HR induction by *Pm3a* and *Pm3f$^{L456P/Y458H}$*. Constructs expressing the AVRPM3$^{A2/F2_G84E}$ and AVRPM3$^{A2/F2_A107V}$ variants were co-infiltrated with those expressing *Pm3a* and *Pm3f$^{L456P/Y458H}$* and HR was assessed after 5 days. No HR was observed for either variant with a single amino acid exchange, in comparison with co-infiltrations of *AvrPm3$^{a2/f2}$*with *Pm3a* or *Pm3f$^{L456P/Y458H}$*. (C) Suppression of the *AvrPm3$^{a2/f2}$-Pm3a/f* mediated HR by *Pm3b* and *Pm3c* in transient assays in *N. benthamiana*. The construct expressing *AvrPm3$^{a2/f2}$* was co-infiltrated with constructs expressing *Pm3a* or *Pm3f$^{L456P/Y458H}$* together with those expressing *Pm3b*, *Pm3c*, or GUS reporter. HR was assessed after 5 days. Suppression of HR was observed with all the combinations including *Pm3b* or *Pm3c* in presence of *AvrPm3$^{a2/f2}$* and *Pm3a*, and with *AvrPm3$^{a2/f2}$* and *Pm3f$^{L456P/Y458H}$*. By contrast, no suppression of HR was observed in a positive control where *Pm3b* and *Pm3c* were replaced by GUS reporter, or when *AvrPm3$^{a2/f2}$*, *Pm3a*, and *Pm3f* were combined. Results in (A), (B), and (C) are consistent across replicates from at least three independent experiments where 4 to 8 leaves were assayed.

Identification and phylogenetic analysis of the *AvrPm3$^{a2/f2}$* effector family

AvrPm3$^{a2/f2}$ encodes a typical *Blumeria graminis* candidate secreted effector protein, which is relatively small (130 residues) and contains a predicted signal peptide, with a cleavage site predicted between position 24 and 25 (SignalP 4.0; Petersen et al., 2011; see Methods), a conserved YxC motif and a second C-terminal cysteine residue (Godfrey et al., 2010; Spanu et al., 2010; Wicker et al., 2013), and no homology to any other characterized fungal protein or functional domain (Figure 4A). One interesting feature of *B.g. tritici* effectors is that they are often organized in clusters of members from the same family (Wicker et al., 2013). To test if this is also the case for the *AvrPm3$^{a2/f2}$*, we screened the genome of the reference 96224 isolate for genes that are closely related using BLAST search for sequence homology. We found an effector family of 24 members including all eight effectors predicted in *locus_2* (see Appendix I: Supplemental Figure 10). We found five members located in sequences genetically close to *locus_2*, mainly in the genetically flanking contigs 33 and 52. Three were found as singletons in genetically unlinked sequences in contigs 81, 143 and 181. Finally, seven members were found in unassembled genomic sequences for which we have no genetic data. Taken together, these results demonstrate that *AvrPm3$^{a2/f2}$* is a member of a larger gene family mainly clustered in *locus_2* and its surrounding genomic regions.

Wicker et al. (2013) have shown that almost 92% of the predicted *B.g. tritici* genes have homologs in *B.g. hordei*. We therefore searched the *B.g. hordei* genome for sequences that were homologous to the *AvrPm3$^{a2/f2}$* effector family, and found a set of 22 closely related effectors. Phylogenetic analysis showed that the sequences arranged into small groups by *formae speciales*, and only Pu_19 from *B.g. tritici* had a direct homologue in *B.g. hordei* (CCU83241.1) (Appendix I: Supplemental Figure 11).We then tested for positive selection as previously described in Yang et al. (2000; 2005). We found that both effector families are under diversifying selection, with the sequence outside of the signal peptide

region accumulating positively selected amino acids. Based on these observations, we inferred that the two effector families probably derived from a few common ancestors before the separation between *B.g. tritici* and *B.g. hordei*. Each family evolved along with host specialization through gene duplication events, leading to family expansion. Finally, accumulation of mutations outside of the signal peptide region resulted in high sequence divergence within the family.

Functional relevance of amino acid polymorphisms between the virulence and avirulence alleles of *AvrPm3$^{a2/f2}$*

Sequence variation in *Avr* effectors is a well-known mechanism that allows pathogens to escape recognition and counter-adapt to new host resistances (Ravensdale et al., 2011; Raffaele and Kamoun, 2012; Huang et al., 2014; Vleeshouwers and Oliver, 2014). To determine the functional relevance of the two encoded sequence polymorphisms between *AvrPm3$^{a2/f2}$*and *avrPm3$^{a2/f2}$* on recognition by *Pm3a* and *Pm3f*, we modified *AvrPm3$^{a2/f2}$* by site directed mutagenesis. Two modified versions carrying either one of the two polymorphisms previously identified between the parental sequences were studied. In the *AvrPm3$^{a2/f}$*$_{G84E}$ version, the glycine at position 84 was replaced by glutamate, while in *AvrPm3$^{a2/f2}$*$_{A107V}$ the alanine at position 107 was replaced by valine. Then, *AvrPm3$^{a2/f2}$*$_{G84E}$ and *AvrPm3$^{a2/f2}$*$_{A107V}$ were transiently co-expressed with *Pm3a* and *Pm3f*$^{L456P/Y458H}$ in *N. benthamiana*. Interestingly, No HR was observer in all infiltrations involving either combinations, indicating that both amino acid polymorphisms are necessary for recognition by PM3A and PM3F, even though the glycine to glutamate substitution might have resulted in a more substantial modification of the protein than the conservative alanine to valine substitution (Figure 6B).

Allelic suppression of the *AvrPm3$^{a2/f2}$-Pm3a/f* interaction

Stirnweis et al. (2014b) previously showed that *Pm3b* was able to suppress an auto-active version of *Pm3f* that mimics downstream cell-death signaling upon *Avr* recognition. Therefore *Pm3b* is referred to as the "Suppressing" allele and *Pm3f* as the "Suppressed" allele. Here, we had the opportunity to address allelic suppression in the presence of the AVR protein. To do so, we co-expressed *AvrPm3$^{a2/f2}$*together with *Pm3f*$^{L456P/Y458H}$ or *Pm3a* in order to induce HR in *N. benthamiana*, and added *Pm3b* or *Pm3c* to test for HR suppression. Both *Pm3b* and *Pm3c* were able to suppress the *Pm3a* and *Pm3f*$^{L456P/Y458H}$ mediated HR in the presence of *AvrPm3$^{a2/f2}$*, demonstrating that the *Pm3b/c* alleles suppress the *Pm3a/f*-mediated resistance signaling also in the presence of *AvrPm3$^{a2/f2}$*and *Avr*-dependent HR (Figure 6C).

Role of locus_1 in the *AvrPm3-Pm3* interaction

Our genetic segregation and genetic mapping data indicate that *locus_1* is involved in all *AvrPm3-Pm3* interactions including the *AvrPm3$^{a2/f2}$-Pm3a/f* one. In our genetic model, avirulence on *Pm3a* is only determined by the genotype at *locus_1* in the 94202 population. According to the 1:1 (A:V) segregation ratio, all progeny receiving the avirulence allele at this locus are avirulent on *Pm3a* (Figure 1A). For *Pm3f*, our genetic model indicates that the avirulence allele is required at both *locus_1* and *locus_2* (*AvrPm3$^{a2/f2}$*) so that a progeny is rendered avirulent on *Pm3f*. However, we have demonstrated that *locus_2* (which encodes for *AvrPm3$^{a2/f2}$*) is sufficient for conferring specificity of recognition by *Pm3a* and *Pm3f* and yet, avirulence on *Pm3a* is only determined by the genotype at *locus_1*. In this context, the role of *locus_1* remains unclear.

In previous work by Parlange et al. (2015), the putative effector gene *Bcg1* was identified as the only gene co-segregating with *AvrPm3f* and *AvrPm3c* in the JIW2 population. The sequence of *Bcg1* is identical in the virulent parents 94202 and JIW2 but polymorphic to the avirulent parent 96224. Using genetic markers derived from high resolution mapping of *AvrPm3f1* in the 94202 population, Parlange et al. (2015) demonstrated that *Bcg1* cogregates with *locus_1*, thus making the avirulence allele of *Bcg1* (*Bcg1^{avr}*) the best candidate for *AvrPm3$^{a1/f1}$*. Using the same experimental conditions to test *AvrPm3$^{a2/f2}$*(OD_{600}=1.2, Avr:R ratio=4:1), we found that none of the parental alleles of *Bcg1* (*Bcg1^{avr}* or *Bcg1^{vir}*) is recognized by *Pm3a* or *Pm3f* $^{L456P/Y458H}$ (see Appendix I: Supplemental Figure 12). This result suggests that *locus_1* is unlikely to encode for a factor that is recognized by *Pm3a* or *Pm3f*. Thus, the genotype of the avirulent parent in this locus is not responsible for specificity towards *Pm3a* and *Pm3f*. Therefore, segregation ratios on *Pm3a* and *Pm3f* are best explained by a model where *locus_1* encodes for a suppressor of the *AvrPm3-Pm3* allelic interactions. In the case of *Pm3a*, the 1:1 segregation is best explained by the presence of a third factor (locus) that is not polymorphic between the parental isolates, and recognized by *Pm3a* only.

Under the hypothesis of an interaction between *Bcg1*, *AvrPm3$^{a2/f2}$*, and *Pm3a/f*, the two effector genes might have coordinated expression kinetics. We isolated RNA from leaf segments of the susceptible wheat cultivar "*Chinese Spring*" inoculated with the avirulent isolate 96224. Samples were collected from three independent biological replicates every 24 hours from 1 to 5 days post inoculation, and relative gene expression over time was assessed by qRT-PCR. Indeed, *Bcg1* and *AvrPm3$^{a2/f2}$* had similar expression kinetics, with a peak at 2 days (Figure 7A), indicating both effectors were up-regulated at the early stages of haustorium formation where interactions with resistance proteins are likely to happen. Also, *AvrPm3$^{a2/f2}$*, which is responsible for the avirulence of 96224 towards *Pm3a/f*, was induced earlier than *Bcg1* (1 dpi) and at a higher level.

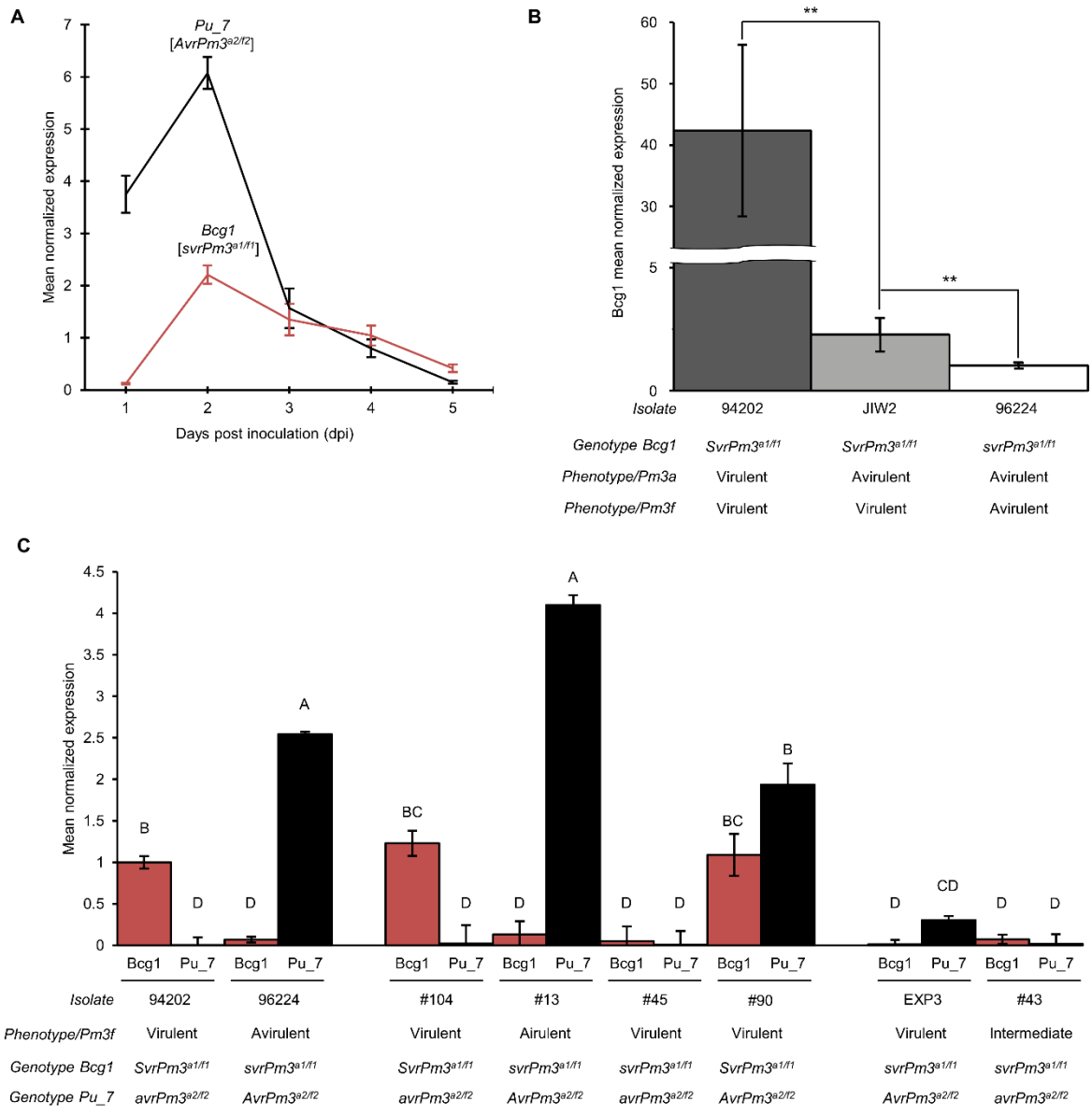


Figure 7. Gene expression analysis of the *SvrPm3*$^{a1/f1}$ candidate suppressor *Bcg1*, and the *AvrPm3*$^{a2/f2}$ avirulence effector *Pu_7*, in parental isolates and F_1 progeny of powdery mildew. (A) The mean normalized expression of *Pu_7* (*AvrPm3*$^{a2/f2}$, black line) and *Bcg1* (*svrPm3*$^{a1/f1}$, red line) in the avirulent parental isolate 96224, 1 to 5 days after infection of the susceptible wheat line "*Chinese Spring*". (B) The mean normalized expression of *Bcg1* at 2 days after infection of the susceptible wheat line "*Chinese Spring*" with the parental isolates 94202, JIW2 and 96224. The phenotypes on *Pm3a* and *Pm3f,* are indicated. The genotypes of *Bcg1* and *Pu_7* are indicated according to the proposed *Avr/Svr* nomenclature (see text). Statistical significance was assessed using the unpaired Student's t-Test. Significance under the threshold $P \leq 0.01$ (**) is indicated. (C) The mean normalized expression of *Bcg1* (red histograms) and *Pu_7* (black histograms) at 2 days after infection. Leaf segments from the susceptible backcrossing line "*Chancellor*" were infected with i) the parental isolates (94202 and 96224), ii) four progeny selected from each allelic combination of the parental genotypes for *Bcg1* and *Pu_7* (#104, #13, #45, #90), iii) the *Pm3f* gain of virulence mutant EXP3 (UV mutagenized 96224), and iv) the *Pm3f* intermediate progeny #43. Progeny are labeled with their identification number as assigned in Figure 2A. The phenotypes on *Pm3f* are indicated. The parental genotypes of *Bcg1* and *Pu_7* are indicated according to the proposed *Avr/Svr* nomenclature (see text). Gene

(Figure 7. Continued from previous page) expression levels were compared using the Tukey-Kramer mean separation procedure (α = 0.05). Significantly different patterns are indicated by different letters on top of the histograms. In all these assays, total RNA was extracted from infected leaf tissue flash frozen at the indicated time points. The values represent the average expression from three independent biological replicates. Standard error of the mean is indicated.

Then, to assess if the expression pattern of *Bcg1*^{avr} and *Bcg1*^{vir} correlate with avirulence/virulence towards *Pm3a* and *Pm3f*, we compared the expression levels in the parental isolates 96224 (*AvrPm3*$^{a1/f1}$, avirulent on *Pm3a* and *Pm3f*), JIW2 (*avrPm3*$^{a1/f1}$, avirulent on *Pm3a* but virulent *Pm3f*) and 94202 (*avrPm3*$^{a1/f1}$, virulent on *Pm3a* and *Pm3f*). RNA samples were produced from three independent biological replicates specifically at 2 days post infection when *Bcg1* expression is highest. Quantitative expression analysis showed that *Bcg1* was much higher expressed in the virulent 94202 compared to 96224 (42.3 fold) and JIW2 (18.9 fold). To a less extent, but still of statistical significance, *Bcg1* expression was also higher in JIW2 compared to 96224 (2.2 fold) (Figure 7B). Assuming that these differences are of biological relevance, they suggest that in addition to sequence variation between the *Bcg1*^{avr} and *Bcg1*^{vir} alleles, expression level polymorphism also correlates with virulence on *Pm3a* and *Pm3f*. The isolate 94202 showed the highest expression and is virulent on both *Pm3* alleles. In the isolate JIW2, which is virulent only on the 'weaker' *Pm3f* allele, *Bcg1* is less expressed than in 94202, but still more than in 96224, which is avirulent on both alleles. Taken together, these data suggest that the *Bcg1*^{vir} is a quantitatively acting virulence factor, encoded by the virulent parent, which can suppress the *AvrPm3*$^{a2/f2}$*-Pm3a/f*-mediated resistance. This suppression possibly depends on the level of gene expression, and ultimately on different amounts of the final protein product required for suppressing the resistance mediated by the 'stronger' PM3A *vs.* the 'weaker' PM3F allele.

To further test the suppressor hypothesis, we first analyzed the genetic association between *Bcg1 / AvrPm3*$^{a2/f2}$ genotypes and virulence on *Pm3a* and *Pm3f* in 133 progeny (out of the 143 for which we had complete phenotype data for both alleles) (see Appendix I: Supplemental Figure 13 and Supplemental Text). According to our genetic model, we expect to find four genotype combinations: *Bcg1*^{avr}*-AvrPm3*$^{a2/f2}$, *Bcg1*^{avr}*-avrPm3*$^{a2/f2}$, *Bcg1*^{vir}*-AvrPm3*$^{a2/f2}$ and *Bcg1*^{vir}*-avrPm3*$^{a2/f2}$. Based on the 1:3 (A:V) segregation on *Pm3f*, we also expect that all progeny receiving the genotype *Bcg1*^{avr}*-AvrPm3*$^{a2/f2}$ are avirulent, while all progeny receiving any of the three other genotypes are virulent. In the 133 progeny we tested, we found perfect association between genotype combination and virulence on *Pm3f* as predicted by our model (see Appendix I: Supplemental Figure 13). In particular, our data showed that all individuals with the genotype *Bcg1*^{vir}*-AvrPm3*$^{a2/f2}$ were indeed all virulent on *Pm3f* and *Pm3a,* which supports the hypothesis that the *AvrPm3*$^{a2/f2}$-mediated avirulence is suppressed by *Bcg1*^{vir} in these progeny.

Using the same agroinfiltration assay previously described in this work, we co-expressed *AvrPm3*$^{a2/f2}$, *Pm3f*$^{L456P/Y458H}$ and *Bcg1*^{vir} in *N. benthamiana* and *N. tabacum*, and assessed for HR suppression in 39 leaves (30 leaves from two biological replicates in *N. tabacum*, and 9 leaves from a third replicate in *N. benthamiana*). HR suppression was assessed 5 days after infiltration relatively to a control where *Bcg1*^{vir} was replaced by a GUS reporter plasmid, and was observed in 9 leaves (23%), and only in the combinations including *Bcg1*^{vir} (see Appendix I: Supplemental Text and Supplemental Figures 14A and 14B). Here, although only partial HR suppression was obtained, we observed no leaves with HR suppression when GUS was delivered instead of *Bcg1*^{vir} (see Appendix I: Supplemental Figure 14B), thus suggesting that *Bcg1*^{vir} can suppress the *AvrPm3*$^{a2/f2}$, *Pm3f*$^{L456P/Y458H}$ -mediated HR.

Then, to test for suppression activity of *Bcg1*^{vir} in wheat we used the same particle bombardment assay as described above for the functional validation of *AvrPm3*$^{a2/f2}$ (see Figure 5B) to co-express *AvrPm3*$^{a2/f2}$, *Pm3a* and *Bcg1*^{vir}. As a control for the *AvrPm3*$^{a2/f2}$-*Pm3a* interaction, we replaced *Bcg1*^{vir} by *Pu_3*, a member of the *AvrPm3*$^{a2/f2}$ effector family that is not recognized by *Pm3a* or *Pm3f*. In a third set of experiments, we have co-expressed *Pm3a*, the empty expression vector and the GUS reporter, which served as a control for GUS accumulation. The number of GUS expressing cells was assessed from two independent bombardment assays. No significant difference was observed between the bombardment involving *Pm3a,* the empty expression vector and the GUS reporter, and the one involving *Bcg1*^{vir} (*Pm3a, AvrPm3*$^{a2/f2}$, *Bcg1*^{vir}). By contrast, a significant difference was observed compared to bombardment with *Pm3a, AvrPm3*$^{a2/f2}$ and *Pu_3* (see Appendix I: Supplemental Figure 14C). As expected, GUS accumulation was compromised in cells co-expressing *AvrPm3*$^{a2/f2}$and *Pm3a* and *Pu_3*. However, in the combination where *Bcg1*^{vir} is added instead of *Pu_3*, the number of GUS expressing cells increases to a level comparable to the *Pm3a*-empty expression vector- GUS reporter control. These results are consistent with the data obtained by transient expression assays in *N. benthamiana* and *N. tabacum*, and support the hypothesis that *Bcg1*^{vir} acts as a suppressor of the *AvrPm3*$^{a2/f2}$-*Pm3a/f* interaction. For clarity and consistency with the genetics, we will hereafter refer to *Bcg1* as a candidate suppressor of avirulence (*Svr*), with the parental alleles *Bcg1*^{avr} *and Bcg1*^{vir} called *svrPm3*$^{a1/f1}$ and *SvrPm3*$^{a1/f1}$, respectively.

Inheritance of parental gene expression levels of *SvrPm3*$^{a1/f1}$ and *AvrPm3*$^{a2/f2}$

To further investigate the role of quantitative differences in gene expression in conferring or suppressing allelic recognition, we analyzed the expression of *SvrPm3*$^{a1/f1}$ and *AvrPm3*$^{a2/f2}$ in the parental isolates (96224, 94202), and in four progeny exemplifying the four different genotype combinations depicted in Figure 1B. We also analyzed gene expression in progeny #43, which shows intermediate avirulence on *Pm3f*. Finally, we included in our analysis the EXP3 mutant (derived from isolate 96224 by UV

mutagenesis), that gained virulence on *Pm3f* and intermediate virulence on *Pm3a* (Parlange et al., 2015). Illumina sequencing of this mutant revealed no polymorphism to the genotype of the avirulent parent in *locus_1* and *locus_2*. Additional resequencing of PCR fragments covering *locus_1* (Parlange et al., 2015) and 1,048 bp upstream and 895 bp downstream of *AvrPm3$^{a2/f2}$* (this work) confirmed that the EXP3 genotype is identical to that of the avirulent parent at these loci (*svrPm3$^{a1/f1}$ -AvrPm3$^{a2/f2}$*).

RNA samples corresponding to three independent infections with the previously cited progeny isolates, and grown on the susceptible wheat line '*Chancellor*', were collected at 2dpi. Relative gene expression was assessed by qRT-PCR as in the previous experiments. In the parental isolates 96224 and 94202 *SvrPm3$^{a1/f1}$* was highly expressed only in the virulent parent while *AvrPm3$^{a2/f2}$* was highly expressed only the avirulent parent. The exact opposite was obtained for *svrPm3$^{a1/f1}$*and *avrPm3$^{a2/f2}$* (Figure 7C). In progeny #104, #13, #45 and #90, we found that relative gene expression levels were best explained by the parental origin of the gene, which indicates that avirulence and suppressor genes are independently regulated. Here, the key observation is in progeny #90 which possesses the recognized *AvrPm3$^{a2/f2}$* and yet it is virulent on *Pm3f.* This result is evidence that the *AvrPm3$^{a2/f2}$-Pm3f* interaction is suppressed in this progeny, which is best explained by the presence and the high expression level of *SvrPm3$^{a1/f1}$*.

For progeny #43, which shows intermediate avirulence on *Pm3f*, gene expression levels were similar to those of the virulent progeny #45. Both isolates have the genotype *svrPm3$^{a1/f1}$-avrPm3$^{a2/f2}$*, and yet they show different virulence levels on *Pm3f*. In the EXP3 mutant, which gained virulence towards *Pm3f* in the 96224 background, the expression pattern was also similar to that of progeny #45. However, EXP3 has the genotype of the avirulent parent for both genes (*svrPm3$^{a1/f1}$*, *AvrPm3$^{a2/f2}$*) which suggests that gain-of-virulence towards *Pm3f* results from down-regulation of the *AvrPm3$^{a2/f2}$* effector gene. Together, these results are evidence for a second layer of control of avirulence/virulence, where additional genetic factors quantitatively determine the final outcome of the *AvrPm3-Pm3* interaction, in addition to the polymorphic amino acids distinguishing the proteins encoded by the *AvrPm3$^{a2/f2}$*and *avrPm3$^{a2/f2}$* alleles.

3.3 Discussion

In this study we aimed at a better understanding of the genetic and molecular basis of avirulence and allelic specificity in the race-specific multi-allelic series of the *Pm3* resistance gene. We used genome-wide approaches to map the genes in the fungus controlling avirulence/virulence towards the *Pm3* alleles, and reported a mechanism of genetic control of avirulence involving three genetically interacting loci that is unique to the *AvrPm3-Pm3* interaction. We cloned the avirulence effector in powdery mildews (*AvrPm3$^{a2/f2}$*) that is specifically recognized by the *Pm3a/f* allele pair. We have

demonstrated that different allelic specificities are controlled by distinct *Avr/avr* loci and mediated by recognition of an allele-specific avirulence effector gene, which is different from our original hypothesis of an allelic series of a single *Avr* gene in the pathogen based on the assumption of a simple co-evolution of a receptor-ligand type of interaction.

Genome-wide approaches to *Avr* mapping in wheat powdery mildew

One limiting factor in genetic mapping is the ability to rapidly derive genotype information from entire genomes of several individuals. For this work, the availability of the genome sequences of the parental isolates used in crosses (Wicker et al., 2013), allowed us to use genome-wide approaches for mapping of the *AvrPm3/avrPm3* loci. We relied on the BAC-based physical map of wheat powdery mildew to derive genome-wide SNPs markers representative of the BAC-contigs. Thus, the wheat powdery mildew physical map can be integrated with the genetic maps we produced, which is a powerful approach for improving genome assembly (Beyer et al., 2007). Furthermore, the use of LTC (Frenkel et al., 2010) significantly improved the quality of the BAC assembly, demonstrating that these algorithms are highly suitable in the very repetitive genome of powdery mildew. Finally, using the KASP high-throughput SNP genotyping technology (He et al., 2014), we rapidly and successfully genotyped the SNP markers on a total of 318 progeny, originating from two crosses. We used 200 markers that are common to both mapping populations which also share the same avirulent parent 96224. This strategy allowed all markers to be anchored to the reference 96224 genome, and circumvents marker ordering conflicts such as local reshuffles and global displacement (Wu et al., 2008; Chen and Yang, 2010), thus allowing robust and genome-wide consensus mapping of the *AvrPm3/avrPm3* loci across two populations.

Identification of *AvrPm3$^{a2/f2}$*

We mapped three loci in the genome where two, *locus_1* and *locus*_3, contained *Avr* factors involved in multiple specificities whereas *locus_2* only contained *AvrPm3f2*. Therefore, we chose to map avirulence specificity towards *Pm3f* in *locus_2*. Mapping the genetic interval containing *AvrPm3^{f2}* was only possible because our phenotyping and genetic data enabled us to distinguish between progeny with the *AvrPm3^{f1}-AvrPm3^{f2}* genotype from the ones with the *AvrPm3^{f1}-avrPm3^{f2}* genotype (see Appendix I: Supplemental Figure 6). The sequence of *locus_2* consisted mainly of transposable elements and a family of putative candidate secreted effector genes, including *AvrPm3$^{a2/f2}$*. In barley powdery mildew, close physical association with transposable elements was reported for AVR_{k1} and AVR_{a10}, although no secreted protein was predicted in this locus (Ridout et al., 2006).

The *AvrPm3*$^{a2/f2}$ gene encodes a putative secreted effector protein of 130 amino-acids with a predicted N terminal secretion signal that is typical of avirulence genes from fungal and oomycete plant pathogens (Dodds et al., 2009). Our findings contrast with the atypical avirulence factors AVR_{k1} and AVR_{a10} isolated from barley powdery mildew, which are classified as atypical avirulence proteins, lacking an N-terminal secretion signal, and associated with transposable elements (Ridout et al., 2006). *AvrPm3*$^{a2/f2}$exhibits allelic variation consisting in two amino acid polymorphisms between the avirulence and virulence versions. Both polymorphisms are individually sufficient for escaping recognition by *Pm3a/f* (Figure 6B), which is consistent with the hypothesis that sequence variation is a common mechanism found in plant pathogens to be beneficial for the evolution of *Avr* effectors towards escaping recognition by the host. Other mechanisms such as pseudogenization or gene loss, might also exist (Ravensdale et al 2011; Raffaele and Kamoun, 2012; Huang et al., 2014; Vleeshouwers and Oliver, 2014).

Consistent with previous studies using auto-active PM3 proteins (Stirnweis et al., 2014a; 2014b), transient expression of *AvrPm3*$^{a2/f2}$with *Pm3a*, *Pm3f* or *Pm3f* $^{L456P/Y458H}$ in *N. benthamiana* led to cell death. In wheat, transient expression of *AvrPm3*$^{a2/f2}$ with *Pm3a* led to a significant reduction in the number of cells accumulating the GUS reporter, as did the auto-active *Pm3a*^{HR} construct, likely because of an activation of *Pm3a* by *AvrPm3*$^{a2/f2}$ resulting in cell-death. Therefore, we propose that the *Pm3* resistance is mediated by hypersensitive cell death response upon *AvrPm3* recognition, that is a common immune response in many *Avr-R* interactions (Howles et al., 2005; Catanzariti et al., 2010; Yoshida et al., 2009; Williams et al., 2011; Cesari et al., 2013). Also, taking into consideration that stronger HR was obtained in *N. benthamiana* when using the *AvrPm3*$^{a2/f2}$ construct without signal peptide, we propose that the *AvrPm3-Pm3* interaction is intracellular, which can take place in the cytosol or the nucleus. This also implies that *AvrPm3*$^{a2/f2}$ is translocated into the plant cell. In the oomycete *Pythium ultimum*, a YxSL[KR] motif was found to be conserved among effectors specifically and highly expressed during infection (Levesque et al., 2010), which somewhat resembles the YxC motif found in powdery mildew effectors, including *AvrPm3*$^{a2/f2}$. One hypothesis is that the YxC motif could be involved in translocation of the effector protein inside the host cell. This could be now addressed in powdery mildews, using the *AvrPm3*$^{a2/f2}$- *Pm3a/f* interaction as a system for such functional studies.

The genetic and molecular basis of specificity

Despite very high sequence similarity between the *Pm3* alleles, including the LRR domain (Bhullar et al., 2009, 2010; Brunner et al., 2010), only *Pm3a* and *Pm3f* are capable of recognizing *AvrPm3*$^{a2/f2}$. This specificity is further demonstrated by allelic suppression of the *Pm3a/f*-mediated signaling, by the *Pm3b/c* alleles, in presence of *AvrPm3*$^{a2/f2}$, thus

substantiating this *Avr* is not recognized by *Pm3b/c*. Previous studies have suggested that polymorphisms in leucine rich repeats 25 to 28 of the PM3 alleles are associated with differences in AVR recognition (Brunner et al., 2010; Sela et al., 2014). This LRR region is highly conserved between the *Pm3a/f* pair of 'weaker' and 'stronger' allele (Brunner et al., 2010) which is consistent with dual recognition of *AvrPm3$^{a2/f2}$*. Interestingly, the same region is also highly conserved between *Pm3b/c* pair of 'weaker' and 'stronger' alleles but divergent from that of the *Pm3a/f* pair (Brunner et al., 2010). Assuming there is a similar situation for *Pm3b/c*, we propose that specificity of recognition towards these two alleles is mediated by a single *AvrPm3b/c* effector gene.

On the pathogen side, we showed that avirulence toward the *Pm3b/c* pair is controlled by the same two loci (*locus_1* and *locus_3*). By analogy to *locus_2*, we propose that *locus_3* controls dual specificity towards *Pm3b* and *Pm3c* in a process similar to the *AvrPm3$^{a2/f2}$-Pm3a/f* interaction. In *locus_1* the only gene present (*Bcg1*) encodes a putative secreted effector (Parlange et al., 2015). At the gene expression level, we found that *Bcg1* is induced at dramatically higher levels in the virulent isolate 94202 as compared to the avirulent isolate 96224. The exact opposite was observed for the *AvrPm3$^{a2/f2}$* effector (*Pu_7*). Our interpretation is that if *Bcg1* was an avirulence factor, one would expect this gene to be induced in the avirulent parent, like *AvrPm3$^{a2/f2}$*. This data is best explained by a model where the virulence allele of *Bcg1* acts as a suppressor of the *AvrPm3-Pm3* interaction. This models also predicts the presence of a third factor that is recognized by *Pm3a* only. The suppressor activity of *Bcg1* was partially demonstrated in transient assays in the tobacco system and in wheat. Although these assays do not provide definitive evidence for suppression, these data is in best agreement with the hypothesis that the virulence allele of *Bcg1*, encoded by *locus_1* in the virulent parents 94202 and JIW2, is a suppressor of the *AvrPm3-Pm3* interaction that we refer to as *SvrPm3$^{a1/f1}$*. We propose that *Bcg1* occurs in two allelic forms differing in expression level: one allele (*svrPm3$^{a1/f1}$*) present in the parental isolate 96224 is expressed at low level, allowing specific recognition and signaling to occur, whereas the allele present in isolates 94202 (*SvrPm3$^{a1/f1}$*) is expressed at high level, suppressing recognition or signaling, and therefore acting as a virulence gene. Here, we predict that in addition to amino-acid polymorphism differentiating the proteins encoded by *svrPm3$^{a1/f1}$* and *SvrPm3$^{a1/f1}$*, polymorphisms in gene expression, which ultimately affects the amount of protein produced, might also play a role in suppression. Therefore, we cannot exclude the possibility that the protein encoded by *svrPm3$^{a1/f1}$* in the avirulent parent has retained a suppressor activity, and that suppression would then also depend on the final amount of protein accumulating in the cell. The fact that the *Bcg1* effector gene family is under diversifying selection (Parlange et al., 2015) suggests possible interaction with a host factor. We propose that *Bcg1* suppresses a wheat factor involved in the *AvrPm3-Pm3* mediated resistance, and acts as a third genetic component of the interaction: *SvrPm3*. Thus, the critical factor defining avirulence *vs.* virulence in the *AvrPm3-Pm3* interaction,

in a specific genetic cross of two mildew isolates, depends on the combination of *Avr/avr* and *Svr/svr* factors. The control of avirulence by a suppression mechanism is also an elegant solution that allows the pathogen to avoid recognition without alteration or deletion of the avirulence gene, thus keeping an active effector. Such a complex mechanism in the pathogen might represent an evolutionary advantage for pathogen fitness. It remains to be determined if this model is confirmed in the study of other isolates and crosses, or if there are even additional levels of control which might be revealed by studying additional isolates.

Additionally, the role of polymorphic amino acids *vs*. gene expression changes in $AvrPm3^{a2/f2}$ remains to be further investigated in additional isolates to fully understand the interactions. Based on expression data, gain-of-virulence towards *Pm3f* of the EXP3 mutant is best explained by the much lower level of expression of $AvrPm3^{a2/f2}$as compared to the avirulent parent. In the stem canker fungus *L. maculans*, RNAi mediated silencing of the *AvrLm6* avirulence effector in the *Rlm6* avirulent isolate v23.1.3, resulted in gain of virulence towards the cognate resistance gene (Fudal et al., 2007). Actually, changes to *Avr* genes can be mediated without modification of the protein sequence through mechanisms such as epiallelic variation, mutations in *cis*-elements, alteration of trans-acting factors, or epigenetic modulation (Gijzen et al., 2014; Bakkeren and Valent, 2014). In the case of the $AvrPm3^{a2/f2}$-*Pm3a/f* interaction, we propose that gain of virulence towards *Pm3f* in the EXP3 mutant background is due to an alteration or loss of a trans-acting factor that regulates the expression of $AvrPm3^{a2/f2}$. In future work, we want to map and clone the mutation underlying the molecular basis of gain of virulence in the EXP3 mutant.

The AvrPm3-Pm3 interaction model

The *Pm3* multi-allelic race-specific resistance is controlled by multiple and genetically interacting loci, contrasting with the control of avirulence reported in many fungal and oomycete plant pathogens based on single genes. This control by distinct loci with an allele-specific cognate effector protein, and a suppressor locus, is quite a different system from the *Avr* allelic variants series of rice blast and Arabidopsis downy mildew, involved in avirulence conferred by the *RPP13* and *Pik* allelic series of *R* genes, respectively (Rose et al., 2004; Hall et al., 2009; Terauchi et al., 2011; Kanzaki et al., 2012). However there are some similarities to the more complex flax *L* gene system involving multiple genetic loci in flax rust that determine avirulence, including an inhibitor locus (Lawrence et al., 1981; Ellis et al., 2007;). Finally, the identification of a suppressor locus for the specific *AvrPm3-Pm3* interaction is reminiscent of the *Avr1* function in *Fusarium oxysporum* which acts both as an avirulence gene as well as a suppressor gene (Houtermann et al. 2008, Gawehns et al., 2014; Ma et al. 2015). It remains to be determined whether $SvrPm3^{a1/f1}$ can

also act as an avirulence gene, possibly for any of the many other *Pm3* genes in the wheat genome.

In contrast to the non-LTR retrotransposon-associated atypical avirulence factors in barley powdery mildew (Avr_{a10} and Avr_{k1}) (Ridout et al., 2006), $AvrPm3^{a2/f2}$is a typical fungal avirulence effector. We have also identified a candidate suppressor of the *Pm3a/f* mediated resistance ($SvrPm3^{a1/f1}$) that is also a typical fungal effector. The finding that the repertoire of typical secreted effectors found in powdery mildews contains possible *Avr* genes and suppressor of race-specific resistance, now allows studying effector biology and race-specific resistance in this agronomically important class of fungal plant pathogens. In addition, the $AvrPm3^{a2/f2}$-*Pm3a/f* - $SvrPm3^{a1/f1}$ interaction is a promising a system for functional studies of conserved effector motifs and protein structure. By identifying the first *Avr* of an allele of *Pm3*, we were also able to combine pathogen and host aspects of the *Pm3* allelic specificity that have been studied for more than a decade. In our model, *Pm3* resistance is mediated by an interaction involving an allele-specific avirulence effector (*Avr)*, a resistance gene allele (*R*), and an allele-unspecific pathogen-encoded suppressor of avirulence (*Svr*), that we propose as the *Avr-R-Svr* model for race-specific resistance based on *Pm3* alleles (Figure 8).

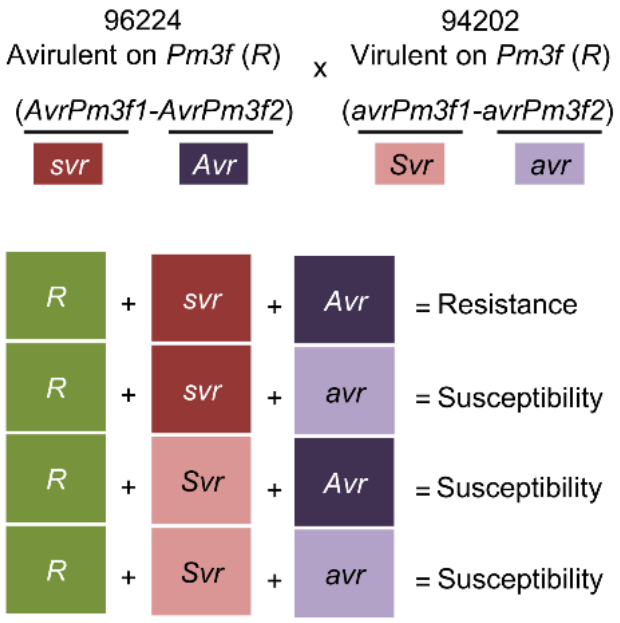


Figure 8. The *Avr/R/Svr* genetic model. We propose that the *Pm3* multi-allelic race-specific resistance follows a three component interaction model involving: i) an allele of the *R* gene; ii) an avirulence effector (*Avr*) specifically recognized by this allele and which can also occur as a virulence form (*avr*) that is not recognized; and iii) a pathogen-encoded suppressor of the *Avr-R* interaction (*Svr*) which can also occur as a non-suppressor form (*svr*). Resistance is mediated only by the *Avr/R/svr* combination where the resistance gene recognizes its cognate avirulence effector in absence of the suppressor. In absence of the cognate *Avr* or in presence of the *Svr* suppressor, the interaction results in susceptibility.

3.4 Materials and Methods

Fungal crosses, plant material, and virulence tests

Bgt isolates were maintained on detached leaves of the susceptible wheat cultivar 'Kanzler' on benzimidazole agar as described in Parlange et al. (2011). The cross between the Swiss isolates 96224 (mating type MAT1-2) and 94202 (mating type MAT1-1) was performed as described in Wicker et al. (2013), and single ascospore cultures were generated and purified twice as described by Brown and Wolfe (1990) producing an F_1 population of 167 individuals. These progeny were entirely or partially phenotyped on wheat differential lines containing only one allele of the *Pm3* resistance gene as follows: 166 on *Pm3a*, 158 on *Pm3b*, 159 on *Pm3c*, 158 on *Pm3d*, 68 on *Pm3e* and 167 on *Pm3f*. Virulence tests were performed as described by Brunner et al. (2010). At least 3 biological replicates were scored at 9 and 12 days post inoculation according to the percentage of leaf surface covered (LC) with pustules. Three distinct phenotypic classes were obtained and scored as follows: LC=60-100%, virulent; LC=10-40%, intermediate; LC=0% avirulent (see Appendix I: Supplemental Figure 1). Of these 167, we have genotyped 154 individuals using the KASP technology. Finally, for mapping multiple *Avr* loci using different subsets of the population, we used 143 progeny out of 154, for which we had complete phenotype data on *Pm3a,b,c,d,f*.

DNA/RNA isolation and construction of plasmid vectors

High molecular weight DNA was extracted as previously described by Wicker et al. (2013) with modifications (see Appendix I: Supplemental Text). RNA samples were extracted from infected leaf material using the Qiagen miRNeasy Mini Kit (Qiagen, Hilden, Germany) according to the manufacturer. RACE-PCR ready cDNA was prepared with the SMARTer RACE cDNA kit (Takara Bio, Shiga, Japan) according to the manufacturer. Full length cDNA was prepared using the SuperScript III RT kit (Life Technologies, Carlsbad, USA) according to the manufacturer. Molecular cloning into gateway compatible entry vector was performed using the pENTR/D-TOPO Cloning Kit (Life Technologies, Carlsbad, USA) according to the manufacturer. Site directed mutagenesis and recombination to the binary vector pIPKb004 were performed as previously described by Stirnweis et al. (2014a). Primers used for experimental gene annotation and molecular cloning are listed in Appendix I: Supplemental table 1.

Quantitative real-time PCR experiments

Quantitative real-time PCR experiments were performed according to the MIQE guidelines (Bustin et al., 2009). RNA samples were extracted from infected leaf material originating from three independent biological replicates, using the Qiagen miRNeasy Mini Kit (Qiagen, Hilden, Germany) according to the manufacturer. qRT-PCR ready cDNA was synthesized from 1.5 µg total RNA using the iScript cDNA synthesis kit (Bio-Rad, Hercules, USA) according to the manufacturer. Glyceraldehyde 3-phosphate dehydrogenase (*Gapdh*) and Actin (*Act1*) genes were used as internal controls. Gene

expression was normalized to that of *Gapdh*. Target and reference genes fragments were amplified using KAPA SYBR FAST qPCR Kit (Kapa Biosystems, Wilmington, USA) and the CFX96™ Real-Time PCR detection system (Bio-Rad, Hercules, USA) according to the manufacturer. Gene expression was analyzed using the CFX Manager software version 3.1 (http://www.bio-rad.com/en-ch/product/cfx-manager-software). qRT-PCR primers for reference and target genes are listed in Appendix I: Supplemental table 2.

Construction of genetic linkage maps

Markers were derived from high confidence SNPs (minimum base coverage threshold ≥ 5x, confirmation from ≥ 90% of the reads) in predicted coding sequences preferentially nearest to the ends of the 250 Fingerprinted BAC contigs of the genome (FPCs, hereafter referred to as contigs, Parlange et al. 2011, Wicker et al. 2013). Contigs that lacked high confidence SNPs in coding sequences, or those that contained no predicted gene, were excluded from the analysis. Large contigs (size >500 kb) were given two markers (left end, LE; and right end, RE), and small contigs were given one marker (middle, MI). The final list of SNP derived markers selected for genotyping is indicated in Appendix I: Supplemental file 1. Progeny were scored for identity to the parental genotypes using the KASP technology (He et al., 2014) at KBioscience (LGC Genomics, Hoddesdon, UK). Monomorphic and null data markers were removed and genetic maps were constructed using the MapDisto software as described in Aluko et al. (2004) (LOD min = 3 and r max = 0.3), Mapmaker as described in Lander et al. (1987), and MultiPoint (www.multiqtl.com). Mapping of *AvrPm3*^{f2}, *AvrPm3*^{b2} and *AvrPm3*^{d2} was done by selecting progeny subsets segregating in a 1:1 ratio (A:V) for these loci, based on overall segregation ratios and phenotype patterns (see Appendix I: Supplemental Figures 2, 3, 4, 6, and Supplemental Text). Map curation and construction of consensus linkage groups carrying the *AvrPm3* loci were done manually by comparing one-by-one marker order in all linkage groups.

Bulk Segregant Analysis

Bulked DNA samples were prepared by mixing an equal ratio of DNA extracted from 23 progeny of the *AvrPm3*^{f1}/*AvrPm3*^{f2} genotype (see Appendix I: Supplemental Figure 6). Five µg of DNA were used for Illumina sequencing of 125 bp paired-ends with 200x coverage (8.7x per progeny). Of the 406.7 million Illumina reads (100 bp on average), 264.7 million reads were aligned to the reference sequence (96224) using CLC Genomics Workbench 6.01 software, resulting in an average read depth of 167.8x. Only SNPs on reads with >20x and <200x coverage where analyzed for polymorphisms to ensure representation of each progeny and to exclude SNPs in transposable elements. The final list of 99,826 SNP positions corresponding to the avirulent 96224 genotype was compiled with a threshold of frequency >90%. SNPs were plotted on the 250 BAC contigs of the genome and association to the 96224 genotype was assessed for each contig manually. We identified 10 contigs genetically associated to the 96224 genotype (see Appendix I:

Supplemental Table 3 and Supplemental Figure 7). Contig-specific CAPS markers were used for achieving high-resolution mapping in *locus_2*.

LTC and BAC sequencing

The *B.g. tritici* BAC library was previously produced from isolate 96224 as described in Parlange et al. (2011). The Linear Topological Contig (LTC) assembly was applied to the genome as described in Frenkel et al. (2010). The LTC assembly resulted in three clusters that were manually combined to form scaffold 10 which contained *locus_2*. LTC cluster visualization and graphical editing was performed using Pajek software (Batagelj and Mrvar, 2003). BACs were selected based on their positions in BAC fingerprint assemblies generated with LTC (Frenkel et al., 2010) and 3D-DNA pools of the library were screened by PCR using genetic and physical markers. Candidate clones were confirmed by PCR and plasmids were extracted using the Qiagen Large-Construct Kit (Qiagen, Hilden, Germany) according to the manufacturer's protocol. BAC insert sizes were estimated by digestion with *Not*I and pulse field gel electrophoresis (PFGE). Selected BAC clones were sequenced using Illumina MiSeq technology (2x 250 bp paired end; GATC Biotech, Constance, Germany). Individual BAC reads were assembled using the CLC Genomics Workbench version 6.0.1 at default settings. Approximately 5-17 contigs were obtained per BAC and were arranged into the most likely linear order based on sequence analysis and information on BAC clone overlaps.

Gene annotation

Annotation of the assembled *locus_2* sequence was performed manually by first identifying repetitive elements (TEs) by BLAST search against an in-house database of *Blumeria* TEs. Genes were identified by BLASTx searches against *Blumeria* proteins (Wicker et al., 2013). Both TEs and genes were manually annotated using dot-plot alignments produced with Dotter (Sonnhammer and Durbin, 1995). Identification of orthologous genes in *B.g. hordei* and *B.g. tritici* was done manually. BLAST search results of gene sequences were analyzed according to the number of detected homologs. Significantly homologous single hits were regarded as single copy orthologs while multiple predicted orthologous hits were disregarded. Based on shared homology at the nucleotide level, Clustalx 2.1 alignments (Larkin et al., 2007) were made by manually selecting the members of the *AvrPm3*$^{a2/f2}$ candidate effector gene family and the homologous family in *B.g. hordei*. Secretion signals were predicted by SignalP V3.0 (D-cutoff values >0.5).

Phylogenetic analysis

All multiple alignments were performed with Muscle 3.8.31 (Edgar 2004). Raxml 8.0.22 (Stamatakis 2014) was used to find the maximum likelihood tree using a GTR + GAMMA model, bootstrap support was computed with 100 replications. The protein alignments were back translated to nucleotide alignments using TranslatorX v1.1 (Abascal et al.,

2010). To test for positive selection we estimated the likelihood of the Maximum Likelihood tree under M8a and M8 model (Yang et al., 2000) with Paml 4.8 (Yang et al., 2007) and subsequently used the likelihood ratio test. To test for positive selection on specific subtrees we used the branch model combined with the site model M2 with ω 2 > 1 on the foreground clade(s) and M2 with ω 2 = ω 1 = 1 on the background (Nielsen and Yang, 1998, Weadick and Chang, 2012).

Transient protein expression assays in N. benthamiana *and* N. tabacum

Transient expression by agroinfiltration in *Nicotiana benthamiana* and *N. tabacum* were performed according to the protocol described by Ma et al. (2012) and Dugdale et al. (2014) respectively, with modifications. Briefly, we added a wash/recovery step where overnight Agrobacterium cultures in Luria Broth (LB) medium containing appropriate antibiotics, were harvested by centrifugation at 3000 x g for 5 min, followed by re-suspension in fresh LB medium without antibiotics and incubation for 30 min to 1 hour at 28°C with 200 rpm shaking. Then, agrobacteria were harvested, re-suspended and diluted in infiltration medium (10 mM morpholineethanesulfonic acid, pH 5.6, 10 mM MgCl 2, 150 µM acetosyringone) to an optical density 600 nm = 1.2, and incubated 3 to 4 hours at 28°C with 200 rpm shaking, to induce virulence. To test for an *AvrPm3-Pm3* interaction, agrobacteria expressing the *AvrPm3* candidates or the *Pm* resistance gene alleles were mixed in 4:1 ratio of *Avr*:*R*. To test for suppression of the *AvrPm3a2/f2-Pm3f*$^{L456P/Y458H}$ interaction by *Bcg1*^{vir} (*SvrPm3*$^{a1/f1}$), agrobacteria expressing each one these genes were mixed in a 3:1:9 ratio of *Avr:R:Svr*, using pIPKb004_GUS as a control. In *N. tabacum*, HR development was assessed 5-7 days after agroinfiltration. In *N. benthamiana*, HR development was revealed 5 days after infiltration by Trypan blue staining as described in Ma et al. (2012).

Transient protein expression assays in wheat

Particle bombardment was performed as described by Brunner et al. (2010) with modifications. To assess for recognition of *AvrPm3*$^{a2/f2}$ by *Pm3a* in wheat, leaf segments were co-bombarded with 1.5 µg of the pUbiGUS reporter plasmid, 1.5µg of the resistance gene expression vector, and 6 µg of the *Avr* expressing vector or the empty vector. Leaves were kept for 48 hours without fungal infection and then GUS stained (Schweizer et al., 1999). For each bombardment assay, six slides containing five randomly selected leaf segments were assessed microscopically for the number of GUS stained cells. The number of GUS expressing cells is representative of the average from four to six bombardment assays. To assess for suppression of the *AvrPm3*$^{a2/f2}$*-Pm3a* interaction by *Bcg1*^{vir} (*SvrPm3*$^{a1/f1}$) in wheat, leaf segments were co-bombarded with 1.5 µg of the pUbiGUS reporter plasmid, 1.5 µg of the resistance gene expression vector, 3 µg of the *Avr* expressing vector, and 3 µg of the vector expressing the candidate suppressor *Bcg1*^{vir}. In the controls, *AvrPm3*$^{a2/f2}$ and *Bcg1*^{vir} were replaced in equivalent amounts with the empty vector and the *Pu_3* effector gene, respectively. Leaves were kept for 48 hours

without fungal infection and then GUS stained (Schweizer et al., 1999). For each bombardment assay, six slides containing five randomly selected leaf segments were assessed microscopically for the number of GUS stained cells. Number of GUS expressing cells is representative of the average from two independent bombardment assays.

Accession numbers

The complete sequence of the *AvrPm3f2* locus is deposited in the EMBL/GenBank data libraries under the accession number KT714072.

Acknowledgements

This work was supported by an 'Advanced Investigator' grant from the 'European Research Council' (ERC-2009-AdG 249996, Durable resistance) and by the 'Swiss National Science Foundation' grant 310030B_144081/1.

Author Contributions

S.B., K.E.M., R.B-D., F.P., T.W. and B.K. designed the research. S.B., K.E.M., R.B-D., F.P., S.R., C.R.P., D.S., and G.H. designed experiments. S.B., K.E.M., R.B-D., F.P., C.R.P., and S.F. performed experiments. S.R., S.O., Z.F., F.M., L.K.S., G.T. and T.W. performed bioinformatics and computational analyses. S.O., Z.F., A.B.K., T.W. and B.K. coordinated *de novo* LTC assembly of the powdery mildew physical map. All authors analyzed data. S.B., K.E.M. and B.K. wrote the paper. S.B., T.W. and B.K. coordinated the study. All authors approved the final article.

CHAPTER 4

Distinct domains of the AVRPM3$^{A2/F2}$ avirulence protein from wheat powdery mildew are involved in immune receptor recognition and putative effector function

Kaitlin Elyse McNally[1], Fabrizio Menardo[1], Linda Lüthi[1], Coraline Rosalie Praz[1], Marion Claudia Müller[1], Lukas Kunz[1], Roi Ben-David[2], Kottakota Chandrasekhar[3], Amos Dinoor[3], Christina Cowger[4,5], Emily Meyers[5], Mingfeng Xue[6,7], Fangsong Zeng[6,7], Shuangjun Gong[6,7,8], Dazhao Yu[†,6,7,8], Salim Bourras[†,1] and Beat Keller[†,1]

[1]Department of Plant and Microbial Biology, University of Zürich, Zollikerstrasse 107, 8008 Zürich, Switzerland

[2]Institute of Plant Science, ARO-Volcani Center, 50250 Bet Dagan, Israel;

[3]Department of Plant Pathology and Microbiology, The Robert H Smith Faculty of Agriculture, Food and Environment, The Hebrew University of Jerusalem, Rehovot 76100, Israel;

[4]United States Department of Agriculture-Agricultural Research Service (USDA-ARS), North Carolina State University, Raleigh 27695, USA;

[5]Department of Plant Pathology, North Carolina State University, Raleigh 27695, USA;

[6]Institute of Plant Protection and Soil Science, Hubei Academy of Agricultural Sciences, 430064 Wuhan, China;

[7]Ministry of Agriculture, Key Laboratory of Integrated Pest Management in Crops in Central China, 430064 Wuhan, China;

[8]College of Life Science, Wuhan University, 430072 Wuhan, China.

Published in New Phytologist, Volume 216, Issue 2, pages 681-695, April 2018

Supplemental Information in Appendix II

Summary

Recognition of the AVRPM3$^{A2/F2}$ avirulence protein from powdery mildew by the wheat PM3A/F immune receptor induces a hypersensitive response after coexpression in *Nicotiana benthamiana*. The molecular determinants of this interaction and how they shape natural *AvrPm3$^{a2/f2}$* allelic diversity is unknown. We sequenced the *AvrPm3$^{a2/f2}$* gene in a worldwide collection of 272 mildew isolates. Using the natural polymorphisms of *AvrPm3$^{a2/f2}$* as well as sequence information from related gene family members, we tested 85 single-residue altered AVRPM3$^{A2/F2}$ variants with PM3A, PM3F, and PM3F$^{L456P/Y458H}$ (modified for improved signaling) in *Nicotiana benthamiana* for effects on recognition. An intact *AvrPm3$^{a2/f2}$* gene was found in all analyzed isolates and the protein variant recognized by PM3A/F occurred globally at high frequencies. Single-residue alterations in AVRPM3$^{A2/F2}$ mostly disrupted, but occasionally enhanced the recognition response by PM3A, PM3F and PM3F$^{L456P/Y458H}$. Residues enhancing hypersensitive responses constituted a protein domain separate from both naturally occurring polymorphisms and positively selected residues of the gene family. These results demonstrate the utility of using gene family sequence diversity to screen residues for their role in recognition. This approach identified a putative interaction surface in AVRPM3$^{A2/F2}$ not polymorphic in natural alleles. We conclude that molecular mechanisms besides recognition drive *AvrPm3$^{a2/f2}$* diversification.

4.1 Introduction

Innate immune responses are crucial for early and efficient detection of infectious pathogens. Plant defense responses against pathogen-caused diseases rely heavily upon immune receptors that perceive infection and elicit a localized cell death response to prevent pathogen growth and proliferation. The immunity conferred by such receptors represents an important source of resistance in agriculture. In particular, intracellular nucleotide-binding, leucine rich repeat receptors (NLRs) mediate resistance by detecting secreted pathogen effectors and activating effector-triggered immune responses (ETI) (Qi & Innes, 2013; Zipfel *et al.*, 2014). ETI is an induced localized cell death, or hypersensitive response (HR), that is especially effective against biotrophic fungal and oomycete pathogens, which need living host tissue to survive.

To complete their life cycle, biotrophic pathogens must establish close associations with the host to acquire nutrients and avoid triggering the immune system (Koeck *et al.*, 2011). To this end, filamentous fungal and oomycete pathogens secrete small effector proteins to inhibit host defense responses or hijack the cellular metabolism (Lo Presti *et al.*, 2015). Fungal and oomycete effectors typically have no predicted homology to known proteins and can be structurally related, but often are not

sequence-related (Win *et al,* 2012; Maqbool *et al*., 2015; Lu *et al*., 2016; Praz *et al*., 2017). Pathogen fitness can depend heavily upon effector function, where the loss or inactivation of important effectors reduces virulence (Huang *et al*., 2006; Bos *et al*., 2010). Recognition of specific effectors by host immune receptors also compromises fitness. Accordingly, effectors that are recognized by resistance (*R*) genes become avirulence factors (*Avrs*) and experience strong selection pressure to evade host recognition.

Interactions between host *R* genes and pathogen *Avrs* are highly specific. This specificity of recognition is considered the main driver of *Avr* allelic diversification. Examining *Avr* natural diversity and the effects of individual polymorphisms on recognition has been highly informative for the study of *R* gene specificity. For example, resistance against *Leptosphaeria maculans* is mediated by dual recognition of the *AvrLm4-7* effector by both *Rlm4* and *Rlm7* resistance genes (Parlange *et al*., 2009). Isolates virulent on *Rlm4* contain *AvrLm4-7* alleles with a single common residue alteration that was functionally validated as specifically disrupting *Rlm4* recognition (Parlange *et al*., 2009; Blondeau *et al*., 2015). Following gain-of-virulence mutations in pathogens, reciprocal evolution of host *R* genes may re-establish recognition and resistance. This was also found in the step-wise evolution of specificity between *Avr-Pik* alleles from *Magnaporthe oryzae* and *Pik* alleles from rice (*Oryzae sativa*) (Kanzaki *et al*., 2012). In contrast to gain-of-virulence on *Rlm4*, *L. maculans* isolates virulent on *Rlm7* primarily contain either deletions or several polymorphisms in *AvrLm4-7* that led to inactivation of the protein (Daverdin *et al*., 2012; Blondeau *et al*., 2015). This demonstrates that in addition to diversification of the AVR protein sequence, successful evasion of recognition can also be mediated by *Avr* deletion or gene inactivation, despite potential fitness consequences (Huang *et al*., 2006; 2010).

Studies of recognition specificity have focused extensively on information from natural polymorphisms in both *Avr* and *R* genes. Variants of the avirulence factor ATR1 from *Hyaloperonospora arabidopsis* are recognized by different alleles of the RPP1 immune receptor from *Arabidopsis* (Krasileva *et al*., 2010). An informed selection of mutations from among 69 polymorphisms in ATR1 variants from eight isolates revealed that distributed recognition surfaces on ATR1 determine the specificity of different RPP1 alleles (Krasileva *et al*., 2010; Chou *et al*., 2011). Similarly, in the study of the recognition specificity of flax (*Linum usitatissimum*) *L5* and *L6* alleles for *AvrL567* variants of the flax rust pathogen (*Melampsora lini*), tests combining several mutations from among 35 naturally polymorphic sites at the *AvrL567* locus in 6 rust strains revealed their additive effect on recognition, suggesting that multiple amino acid contact points also determine specificity in this interaction (Ravensdale *et al*., 2012). However, only four AVR-Pik variants were identified in a worldwide screen of 39 *M. oryzae* isolates (Kanzaki *et al*., 2012), demonstrating that even in larger worldwide diversity screens

limited sequence diversity is often observed in fungal *Avr* genes. Therefore, *in vitro* created sequence diversity would be useful for determining the basis of recognition specificity for fungal *Avr-R* interactions.

Active *R* genes are known to coevolve with their cognate pathogen *Avrs* under strong diversifying selection, or more rarely by balancing selection. They often form larger gene families and complex clusters of gene paralogs, but rarely form true allelic series conferring race-specificity (Ellis *et al.*, 1999; Rose *et al.*, 2004; Seeholzer *et al.*, 2010; Kanzaki *et al.*, 2012). Early descriptions of the cognate *Avrs* of multiallelic race-specific *R* genes indicated that allelic variants recognize naturally occurring AVR variants in the pathogen (Dodds *et al.*, 2006; Krasileva *et al.*, 2010; Kanzaki *et al.*, 2012). However, recent advances in the study of multiallelic *R* genes conferring resistance to the cereal powdery mildews (*Blumeria graminis* ff. spp.), specifically the *Pm3* gene in wheat (*Triticum aestivum*) and the *Mla* gene in barley (*Hordeum vulgare*), suggest that distinct resistance alleles recognize highly sequence-diverse effectors (Bourras *et al.*, 2015; Lu *et al.*, 2016).

Powdery mildew of wheat is caused by the obligate biotrophic fungal pathogen, *Blumeria graminis* f. sp. *tritici* (*B.g. tritici*). *B.g. tritici* followed the spread of wheat cultivation to all major wheat growing regions worldwide. The multi-allelic *Pm3* gene from wheat encodes a coiled-coiled (CC) -NLR protein that confers race-specific resistance against *B.g. tritici*. The seventeen functionally distinct alleles of *Pm3* in the hexaploid bread wheat gene pool share particularly high sequence identity (>97%) (Yahiaoui *et al.*, 2004; Srichumpa *et al.*, 2005; Bhullar *et al.*, 2010). The *Pm3a* and *Pm3f* alleles share overlapping recognition spectra toward powdery mildew races, where the spectrum of races recognized by *Pm3a* includes those recognized by *Pm3f* (Brunner *et al.*, 2010). This overlap is the result of their shared specificity for the $AvrPm3^{a2/f2}$ gene encoded by *B.g. tritici*. The $AvrPm3^{a2/f2}$ gene belongs to a family of 24 sequence-divergent but structurally-related secreted effectors (Bourras *et al.*, 2015). In transient expression assays in *Nicotiana benthamiana*, it was demonstrated that $AVRPM3^{A2/F2}$ is recognized specifically by the *Pm3a* and *f* alleles. The hypersensitive response elicited by PM3A recognition is much stronger than the PM3F allele; however, it was demonstrated that two substitutions (L456P, Y458H) in the ARC2 subdomain of the NBS of PM3F are sufficient to enhance this response to levels comparable to PM3A (Stirnweis *et al.*, 2014). In contrast to gene-for-gene interactions (Flor, 1971), segregating phenotypes among the progeny of a genetic cross of isolates hinted at the involvement of a second pathogen-encoded factor, the suppressor-of-recognition $SvrPm3^{a1/f1}$, which was cloned and functionally validated. Expression analyses of both fungal effector genes suggested that the $SvrPm3^{a1/f1}$ suppressor acts quantitatively to suppress $AvrPm3^{a2/f2}$ recognition mediated by PM3A/F (Bourras *et al.*, 2015). Altogether, these findings led to the development of the *Avr-R-Svr* model of interaction in the wheat –

powdery mildew pathosystem, where *R* specificity for *Avr* recognition is modified by the action of a *Svr* (Bourras *et al.*, 2016).

In this study we examined the basis of *Pm3a/f* recognition specificity using natural sequence diversity of *AvrPm3$^{a2/f2}$* from a worldwide collection of *B.g. tritici* and *B.g. triticale*. We found that *AvrPm3$^{a2/f2}$* is present in all isolates and shows limited natural diversity worldwide. We provide further evidence of the role of the suppressor *SvrPm3$^{a2/f2}$* in increasing the virulence of isolates from Europe that express the active *AvrPm3$^{a2/f2}$*. Using the sequence diversity from the structurally related *AvrPm3$^{a2/f2}$* effector family, we identified a region of AVRPM3$^{A2/F2}$ where mutations strongly influence recognition and specificity. This putative interaction domain does not overlap with residues under positive selection in the effector family or residues polymorphic in the natural isolates. In light of our results, we conclude that using sequence diversity from a related gene family is informative for studies of recognition specificity, and propose that for the *AvrPm3$^{a2/f2}$-Pm3a/f* interaction, selection pressure from recognition is not the primary source of *AvrPm3$^{a2/f2}$* allelic diversity.

4.2 Results

A worldwide survey reveals ubiquitous presence and limited sequence diversity of *AvrPm3$^{a2/f2}$*

To study the natural sequence diversity of AVRPM3$^{A2/F2}$ and characterize the variants eliciting recognition by *Pm3a/f*, we sequenced the complete *AvrPm3$^{a2/f2}$* gene in an unprecedented worldwide collection of 272 isolates with diverse genetic backgrounds and a balanced representation of global geographic origins (Supporting Information Table S3). The 251 *B.g. tritici* and 22 *B.g. triticale* (powdery mildew of triticale) isolates form six geographically distinct populations: United States, 56 isolates from 13 states; China, 101 isolates from 12 provinces (Zeng *et al.*, 2014); Europe, 51 isolates from 6 countries; Israel, 61 isolates (Ben-David *et al.*, 2016); Australia, 2 isolates; Japan, 2 isolates. All sampled isolates within this collection encoded a complete *AvrPm3$^{a2/f2}$* gene, with no case of nonsense mutation, truncation, or complete gene deletion. While it is impossible to identify identical gene duplications by sequencing of PCR products, our whole gene sequencing data yielded no evidence of mixed gene sequences that would indicate multiple divergent *AvrPm3$^{a2/f2}$* sequences encoded by a single isolate. In addition, *de novo* analysis of 41 available whole genome assemblies using iterative BLAST searches of the *AvrPm3$^{a2/f2}$* gene sequence revealed no evidence of duplications or multiple haplotypes encoded in a single isolate. We identified 11 novel *AvrPm3$^{a2/f2}$* haplotypes in addition to the two previously reported ones (Bourras *et al.*, 2015). The 13 haplotypes share 95-99% identity, and 12 encode unique protein sequences (labelled A-M, Table 1). All novel protein encoding haplotypes were produced by site-directed mutagenesis or gene synthesis and tested in transient assays in *Nicotiana benthamiana*

for recognition by PM3A, PM3F, and PM3F$^{L456P/Y458H}$ as described in Bourras *et al.* (2015). We used the previously validated AVRPM3$^{A2/F2}$-A as positive control for HR induction upon recognition by PM3A and PM3F$^{L456P/Y458H}$. We observed by visual inspection and fluorescence imaging (Praz *et al.*, 2017) of the assayed *N. benthamiana* leaves at 5 days post infiltration that none of the new variants induced HR in the presence of PM3A or PM3F$^{L456P/Y458H}$, suggesting they encode inactive AVR proteins (Supporting Information Fig. S1). These results indicate that AVRPM3$^{A2/F2}$-A (Bourras *et al.*, 2015) is the only active AVR variant in natural isolates.

Table 1. Disrupting recognition by polymorphisms in the amino acid sequences of the AVRPM3$^{A2/F2}$ variants from wheat powdery mildew.

	No.	21	24	25	26	27	31	38	52	66	69	80	86	89	91	93	95	109	119	122	123
AVRPM3$^{A2/F2}$-A[a]	**122**	**A**	**S**	**G**	**-**	**-**	**N**	**H**	**E**	**N**	**R**	**N**	**G**	**N**	**K**	**E**	**F**	**A**	**Y**	**T**	**E**
AVRPM3$^{A2/F2}$-C	28	.	.	.	-	-	.	.	.	.	.	.	.	.	.	.	L	.	.	.	.
AVRPM3$^{A2/F2}$-E	16	.	.	.	-	-	.	.	.	.	.	.	.	.	.	D	.	.	.	.	.
AVRPM3$^{A2/F2}$-F[b]	15	.	.	.	-	-	.	.	.	K	S	.	E	.	.	.	.	.	.	.	.
AVRPM3$^{A2/F2}$-D	24	.	.	.	-	-	.	.	.	K	.	S	.	.	T	K	.	.	.	.	.
AVRPM3$^{A2/F2}$-B	52	.	.	.	-	-	.	Q	.	.	.	.	E	.	.	.	.	.	.	.	D
AVRPM3$^{A2/F2}$-G	4	V	N	S	P	V	E	Q	.	K	.	.	E	Y	.	.	.	V	H	.	.
AVRPM3$^{A2/F2}$-M	1	.	.	.	-	-	E	Q	N	.	.	.	E	.	.	.	.	.	.	.	D
AVRPM3$^{A2/F2}$-K	2	.	.	.	-	-	.	Q	.	.	.	.	E	.	.	.	.	.	.	.	.
AVRPM3$^{A2/F2}$-H	4	.	.	.	-	-	.	.	.	.	.	.	E	.	.	.	.	.	.	.	D
AVRPM3$^{A2/F2}$-J	3	.	.	.	-	-	.	.	.	.	.	.	E	.	.	.	.	.	.	R	D
AVRPM3$^{A2/F2}$-L[a]	1	.	.	.	-	-	.	.	.	.	.	.	E	.	.	.	.	V	.	.	.
Total	272																				

Polymorphisms compared to the avirulent variant are depicted using the one-letter code for amino acids. Residue numbering includes the signal peptide. Number of isolates encoding each variant in the global collection of 272 isolates is given (No.). Residues that individually disrupt recognition by PM3A and PM3F$^{L456P/Y458H}$ are indicated with black boxes. Untested residues are in grey boxes. Dashes indicate gaps and dots indicate identical residues.

[a]Bourras *et al.*, 2015

[b]Variant found in isolates collected on *T. dicoccoides* (wild emmer)

Genotyping the mildew collection revealed that the *AvrPm3$^{a2/f2}$-A* haplotype is the most common worldwide (122 isolates, 45%), and is present in all six geographic populations at varying frequencies (15-88%, excluding Australia and Japan; Fig. 1a). We obtained phenotypic data on both *Pm3a* and *Pm3f* wheat for 166 isolates in the collection (Table 2). Comparing *AvrPm3$^{a2/f2}$* genotypes and isolate phenotypes on *Pm3a* and *Pm3f*, we observed that, of the isolates encoding any of the inactive AVRPM3$^{A2/F2}$ variants B-M, 72% (65/90) were virulent on *Pm3a*, while 78% (70/90) were virulent on *Pm3f* (Table 2). However, all isolates encoding variant 'F' (also not recognized in our transient assay) were consistently avirulent on *Pm3a* and *Pm3f*, with one exception. We

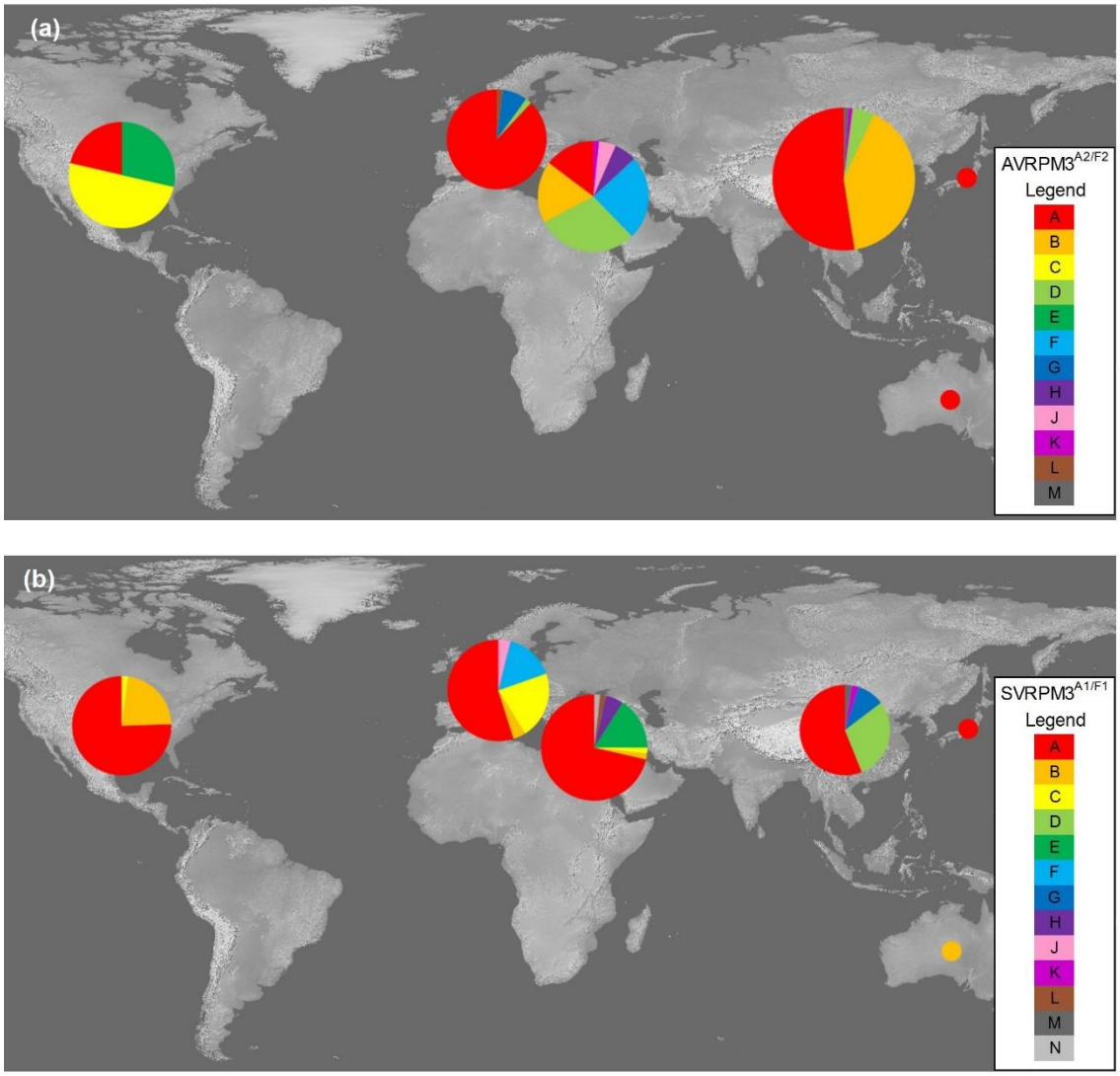


Figure 1. Geographic distribution of the AVRPM3$^{A2/F2}$ and SVRPM3$^{A1/F1}$ variants. (a) Geographic distribution of the 12 AVRPM3$^{A2/F2}$ variants (A-M) encoded by *Blumeria graminis* f. sp. *tritici* and *triticale* isolates from the US, Europe, Israel, China, Japan, and Australia. (b) Geographic distribution of the SVRPM3$^{A1/F1}$ variants (A-N) encoded by *Blumeria graminis* f. sp. *tritici* and *triticale* isolates from the US, Europe, Israel, China, Japan, and Australia. Areas of the circles are proportional to the number of isolates from that region in the tested collection.

also tested the hypothesis that variant 'F' is recognized by other *Pm3* alleles using the same assays in *N. benthamiana*. We found no evidence of interaction with *Pm3b*, *Pm3c*, *Pm3d*, and *Pm3d* alleles, or with the mildew resistance gene *Pm8* (Supporting Information Fig. S2). Evidence of a third factor (*AvrPm3^{a3}*) recognized by *Pm3a* was previously found in a genetic cross of two mildew isolates segregating on *Pm3a* and *Pm3f* (Bourras *et al.*, 2015). There is also genetic evidence from one UV mutagenized mildew strain (EXP3) that a fourth factor is regulating the *AvrPm3$^{a2/f2}$* gene expression (Parlange et al., 2015; Bourras et al. 2015). Therefore, it is likely that additional *Avr* factors or *Avr* modifiers polymorphic in global populations could explain the

avirulence of isolates harboring AVRPM3A2/F2-F on *Pm3a/f*. Another possibility is that *AvrPm3$^{a2/f2}$* is recognized by another NLR in wheat. While evidence from our genetic mapping populations indicates that *AvrPm3$^{a2/f2}$* is not recognized by *Pm1a, Pm2, Pm3b-e, Pm3g, Pm4a, Pm4b, Pm5a, Pm8,* and *Pm17,* we cannot exclude possible recognition by NLRs other than the ones we have tested (Parlange et al., 2015; Bourras et al. 2015; Praz et al. 2017; Supporting Information Notes S1, Tables S4-S6). In isolates encoding the recognized *AvrPm3$^{a2/f2}$-A* haplotype, we observed that 71% (54/76) were avirulent on *Pm3a* wheat, and 17 of these (22%) were also avirulent on *Pm3f* wheat (Table 2). Together, these results suggest that additional pathogen-encoded factors besides *AvrPm3$^{a2/f2}$* are contributing to recognition by *Pm3a/f* in natural mildew populations.

Table 2. Summary of AVRPM3$^{A2/F2}$ variants and phenotypes on *Pm3a* and *Pm3f* wheat observed in natural isolates.

Number of isolates	AVRPM3$^{A2/F2}$ variant	*Pm3a* wheat	*Pm3f* wheat
17	A	A	A
5	A	A	I
32	A	A	V
6	A	I	V
16	A	V	V
17	B	V	V
10	C	V	V
17	D	V	V
8	E	V	V
14	F	A	A
3	G	V	V
3	H	V	V
3	J	V	V
2	B	A	V
2	C	I	I
2	D	A	A
1	D	A	V
1	D	V	A
1	F	A	V
1	G	A	V
1	H	A	V
1	K	A	A
1	K	V	V
1	L	V	V
1	M	V	V

AVRPM3$^{A2/F2}$ variants (A-M) found in natural *Blumeria graminis* f. sp. *tritici* and *triticale* isolates are listed, together with the observed phenotypes on *Pm3a* and *Pm3f* wheat. Variant A (red) is recognized by PM3A and PM3F$^{L456P/Y458H}$ in functional assays. Phenotypes (avirulent, A (green); intermediate, I (yellow); virulent, V (pink)) are colored for ease of reading.

To summarize, our results reveal limited natural diversity of *AvrPm3$^{a2/f2}$*, and the prevalence of the AVRPM3$^{A2/F2}$-A protein variant, suggesting an important role of *avr/AvrPm3$^{a2/f2}$* and particularly the active AVRPM3$^{A2/F2}$-A allele in pathogen fitness. Additionally, our phenotype data suggest that other factors are involved in the interaction. For example, an as yet uncharacterized *AvrPm3^{a3}* avirulence gene might be rendering isolates without the active AVRPM3$^{A2/F2}$-A variant avirulent on *Pm3a*, while the suppressor-of-recognition, *SvrPm3$^{a1/f1}$*, is likely increasing the virulence in isolates encoding the active AVRPM3$^{A2/F2}$-A variant.

Virulence in isolates encoding the recognized AVRPM3$^{A2/F2}$-A is associated with the SVRPM3$^{A1/F1}$ suppressor

Our phenotype data indicate that of the 54 of the 76 isolates encoding the AVRPM3$^{A2/F2}$-A variant that are virulent on *Pm3f*, 16 are also virulent on *Pm3a* despite the presence of the recognized *AvrPm3$^{a2/f2}$* allele (Table 2). Previously a suppressor-of-recognition, *SvrPm3^{a1lf1}*, was functionally validated as quantitatively suppressing the recognition of *AvrPm3$^{a2/f2}$* by *Pm3a/f* in a genetic cross of two European mildew isolates (Bourras *et al.*, 2015). Polymorphic presence of the suppressor might cause the high variability of virulence in isolates encoding the active AVRPM3$^{A2/F2}$-A variant on *Pm3a* and *Pm3f* wheat. To determine whether the presence of the active suppressor is associated with virulence, we obtained sequences of the *SvrPm3$^{a1/f1}$* gene from 201 isolates from the worldwide collection (for which genomic DNA or genome sequences were available), and compared the *Svr* genotypes and phenotypes of the subset of 76 isolates encoding the active *AvrPm3$^{a2/f2}$*. As with *AvrPm3$^{a2/f2}$*, we found a complete *SvrPm3$^{a1/f1}$* gene in all isolates, suggesting that it also contributes to pathogen fitness.

Table 3. Amino acid sequences of the SVRPM3$^{A1/F1}$ variants in wheat powdery mildew.

	No.	22	31	32	70	71	102	118	125	127
SVRPM3$^{A1/F1}$-A^{a}	**130**	**I**	**K**	**P**	**A**	**R**	**L**	**M**	**Y**	**T**
SVRPM3$^{A1/F1}$-B^{a}	16	.	.	.	.	H	.	.	.	.
SVRPM3$^{A1/F1}$-C^{a}	13	.	.	.	.	H	P	.	.	.
SVRPM3$^{A1/F1}$-N^{a}	1	.	.	H	.	H	P	.	.	.
SVRPM3$^{A1/F1}$-F^{a}	8	.	.	.	.	H	P	.	.	K
SVRPM3$^{A1/F1}$-J	2	.	N	.	.	H	P	.	.	.
SVRPM3$^{A1/F1}$-H	3	T	.	.	.	H	P	.	.	.
SVRPM3$^{A1/F1}$-E^{ab}	9	.	.	.	S	H	S	.	.	.
SVRPM3$^{A1/F1}$-D	12	.	.	.	.	H	F	.	.	.
SVRPM3$^{A1/F1}$-M	1	.	.	.	.	H	F	V	.	.
SVRPM3$^{A1/F1}$-K	1	.	.	.	.	.	F	.	.	.
SVRPM3$^{A1/F1}$-G	4	.	.	.	.	.	.	V	.	.
SVRPM3$^{A1/F1}$-L	1	.	.	.	.	.	.	.	C	.
Total	201									

Polymorphisms compared to the known active suppressor variant are depicted using the one-letter code for amino acids. Residue numbering includes the signal peptide. Dashes indicate gaps and dots indicate identical residues.

[a]Parlange *et al.*, 2015

[b]Variant found in isolates collected on wild emmer (*Triticum dicoccoides*)

We identified seven new haplotypes in addition to the six haplotypes previously described (Parlange *et al.*, 2015), for a total of 13 unique but highly similar protein variants (98-99% sequence identity; labelled A-N, Table 3). The previously reported

Table 4. Summary of unique AVRPM3$^{A2/F2}$ and SVRPM3$^{A1/F1}$ variant combinations and isolate phenotypes on *Pm3a* and *Pm3f* observed in natural isolates.

Number of isolates	AVRPM3$^{A2/F2}$ variant	SVRPM3$^{A1/F1}$ variant	*Pm3a* wheat	*Pm3f* wheat
3	A	A	A	I
19	A	A	A	V
5	A	A	I	V
14	A	A	V	V
8	A	B	A	V
9	A	C	A	A
5	A	F	A	A
10	B	A	V	V
9	C	A	V	V
15	D	A	V	V
5	E	A	V	V
3	J	A	V	V
6	B	D	V	V
9	F	E	A	A
1	A	A	A	A
2	A	J	A	A
2	A	D	A	V
2	A	F	A	I
1	A	F	A	V
1	A	G	A	V
1	A	G	I	V
1	A	C	V	V
2	C	A	I	I
1	K	A	A	A
2	D	A	A	A
1	D	A	A	V
1	D	A	V	A
1	H	A	A	V
2	H	A	V	V
1	G	A	V	V
1	L	A	V	V
1	B	K	A	V
1	D	L	V	V
1	D	N	V	V
1	E	B	V	V
2	F	H	A	A
1	F	B	A	A
1	F	C	A	A
1	F	H	A	V
1	G	C	A	V
2	G	B	V	V

active suppressor variant, SVRPM3$^{A1/F1}$-A, is also the most predominant (130 isolates) globally (Fig. 1b).

For 155 isolates we obtained both *AvrPm3$^{a2/f2}$* and *SvrPm3$^{a1/f1}$* sequences, as well as phenotypic data on *Pm3a* and *Pm3f* wheat (Table 4). Of these isolates, 96 encode the active suppressor and among these, 91 are virulent on *Pm3f* and 68 are virulent on *Pm3a*. Strikingly, with only one exception, all isolates encoding the recognized AVRPM3$^{A2/F2}$-A variant and the active suppressor variant SVRPM3$^{A1/F1}$-A (42 isolates) show virulence or intermediate virulence on *Pm3f*, with only one exception. By contrast, all isolates encoding AVRPM3$^{A2/F2}$-A and suppressor variants other than SVRPM3$^{A1/F1}$-A (32 isolates) are avirulent on *Pm3a*, with two exceptions (Table 4), suggesting most SVR variants other than 'A' do not function as active suppressors. The SVRPM3$^{A1/F1}$-B variant might have low suppressor activity, as we found eight isolates encoding AVRPM3$^{A2/F2}$-A and SVRPM3$^{A1/F1}$-B that were virulent on *Pm3f* wheat.

Unique AVR-SVR variant combinations from *Blumeria graminis* f. sp. *tritici* and *triticale* isolates and their observed phenotypes on *Pm3a* and *Pm3f* are shown. AVRPM3$^{A2/F2}$ variants (A-J) and SVRPM3$^{A1/F1}$ variants (A-F, H, and J) are indicated. AVRPM3$^{A2/F2}$ 'A' is the only recognized AVR variant (red), and SVRPM3$^{A1/F1}$ 'A' is the known active suppressor variant (grey, Bourras *et al.*, 2015). Phenotypes (avirulent, A (green); intermediate, I (yellow); virulent, V (pink)) are colored for ease of reading. Variant-phenotype combinations observed in less than three isolates are listed separated for clarity.

These results indicate that *Pm3f* resistance can be suppressed by SVRPM3$^{A1/F1}$-A independently of the mildew isolate genetic background.

Most of the phenotypic variability of isolates encoding both the active AVRPM3$^{A2/F2}$ and SVRPM3$^{A1/F1}$ is observed in the European population where the suppressor was originally described. It is also the population where the combination of both active AVRPM3$^{A2/F2}$ and SVRPM3$^{A1/F1}$ is most frequent (24/50 isolates, 48%, Supporting Information Fig. S3). Within the European population we identified 12 isolates that combine both active variants but are avirulent on *Pm3a* wheat (Supporting Information Table S3). To test if this phenotype can be explained by low expression of the suppressor gene, we compared gene expression levels of *AvrPm3$^{a2/f2}$-A* and *SvrPm3$^{a1/f1}$-A* in two *Pm3a* virulent European *B.g. tritici* isolates (7004 and 07296), and two *Pm3a* avirulent European *B.g. tritici* isolates (07237 and JIW2) at 2 days post infection (Fig. 2). The isolates virulent on *Pm3a* wheat (7004 and 07296) showed similar (07296) or higher (7004) expression levels of *SvrPm3$^{a1/f1}$-A* as compared to *AvrPm3$^{a2/f2}$-A*, consistent with the hypothesis that virulence in these isolates is mediated by relatively higher *Svr* expression. In the *Pm3a* avirulent isolates we found that *SvrPm3$^{a1/f1}$-A* was expressed at significantly lower levels than *AvrPm3$^{a2/f2}$-A*, and also as compared to the expression of the suppressor in the *Pm3a* virulent isolates, 7004 and 07296 (Fig. 2, Supporting Information Table S7). These results substantiate our hypothesis that avirulence on *Pm3a*, in the subset of 12 European isolates combining active AVRPM3$^{A2/F2}$ and SVRPM3$^{A1/F1}$ variants, can be explained by lower expression levels of *SvrPm3$^{a1/f1}$-A*. In addition we compared gene expression levels of *AvrPm3$^{a2/f2}$-A* and *SvrPm3$^{a1/f1}$-A* in 6 European *B.g. triticale* isolates, 3 Israeli *B.g. tritici* isolates, and 1 Japanese *B.g. tritici* isolate with various phenotypes on *Pm3a* and *Pm3f* (Fig. 2). We observe some association between *SvrPm3$^{a1/f1}$-A* expression relative to *AvrPm3$^{a2/f2}$-A* expression and virulence in the *B.g. triticale* isolates, but no association in isolates from Israel and Japan. These data suggest that the quantitative action of a suppressor might be most observable in the European population, and that other factors known to be involved in the interaction (e.g. *AvrPm3^{a3}*) might prevent observable associations between phenotype and *SvrPm3$^{a1/f1}$-A* expression in other populations. In an analysis of the expression of *AvrPm3$^{a2/f2}$-A* with other *SvrPm3$^{a1/f1}$* haplotypes (*B, C, D, F, G,* and *J*) we did not observe any similar association between expression and virulence, although expression of these variants was overall lower than *SvrPm3$^{a1/f1}$-A,* and appeared to be more consistent within isolates and less polymorphic between isolates encoding the same haplotype (Supporting Information Fig. S4). Taken together, our results demonstrate that recognition of *AvrPm3$^{a2/f2}$-A* encoding isolates by *Pm3a/f* wheat is strongly associated to the presence and expression level of *SvrPm3$^{a1/f1}$-A* in European isolates, thus providing additional evidence for the *Avr-R-Svr* genetic model at the population level.

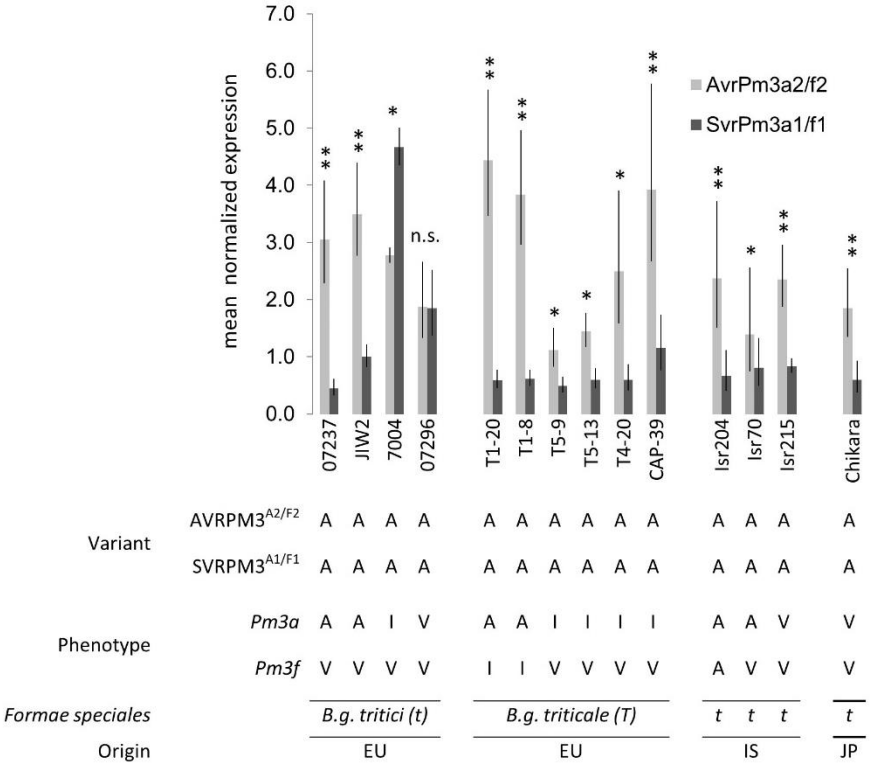


Figure 2. Expression analysis of isolates encoding the AVRPM3$^{A2/F2}$-A and SVRPM3$^{A1/F1}$-A variant combination which give different phenotypes on *Pm3a* and *Pm3f* wheat. The mean normalized expression of *AvrPm3$^{a2/f2}$* (light grey histograms) and *SvrPm3$^{a1/f1}$* (dark grey histograms) at 2 days post infection. Leaf segments from the susceptible recurrent parent line '*Chancellor*' were infected with *B.g. tritici (t)* and *B.g. triticale (T)* isolates from diverse geographic origins (Europe, EU; Israel, IS; Japan, JP) that contain the same *Avr-Svr* genotype, but have different virulences (avirulent, A; intermediate, I; virulent, V) on *Pm3a* and *Pm3f* wheat. The results of a t-test comparing *AvrPm3$^{a2/f2}$* and *SvrPm3$^{a1/f1}$* expression is given by * and **, indicating p-values <0.05 and 0.01, respectively (ns, not significant). The error bars indicate the standard error of the mean (SEM).

Single residue changes in AVRPM3$^{A2/F2}$ are sufficient to abolish, enhance, or alter recognition by PM3A, PM3F and PM3F$^{L456P/Y458H}$

In the study of recognition specificity of *L5* and *L6* for *AvrL567*, Ravensdale and colleagues (2012) chose a subset of 4 residues among 35 polymorphic sites associated with differences in specificity among *AvrLm567* alleles and tested their individual contribution toward specificity. More recently, Maqbool and colleagues (2015) used protein structural information to design four mutations in AVR-PikD they predicted would disturb binding to the PikP cognate NLR. Since neither excessive natural polymorphisms nor protein structural information were available for AVRPM3$^{A2/F2}$, our strategy was to combine information from the 17 natural polymorphisms with an analysis of sequence diversity from the previously described *AvrPm3$^{a2/f2}$* effector family (Bourras *et al.*, 2015). AVRPM3$^{A2/F2}$ belongs to a family of 24 small effectors (90-111 residues in the mature peptide) sharing very little sequence similarity but an overall structural homology (Bourras *et al.*, 2016). We wanted to select single residue

alterations across the entire AVRPM3$^{A2/F2}$ protein to identify amino acid residues affecting recognition specificity by *Pm3a* and *Pm3f*. We assumed that highly conserved residues among the 24 family members were important for structure and that substitutions in AVRPM3$^{A2/F2}$ with residues conserved in many other family members would be structurally conservative. Following this strategy for large scale mutagenesis that conserved protein structure, we individually altered residues in AVRPM3$^{A2/F2}$ toward residues shared by at least 6 (25%) family members (Fig. 3). In addition, the three most highly conserved residues in the family, Y35, C37, and C118 were individually changed to glycine to test their role in recognition, and mutations at residues 102, 128, and 132 were included for even coverage using either the next most frequent residue, or the residue encoded by the close relative, *Pu_23*.

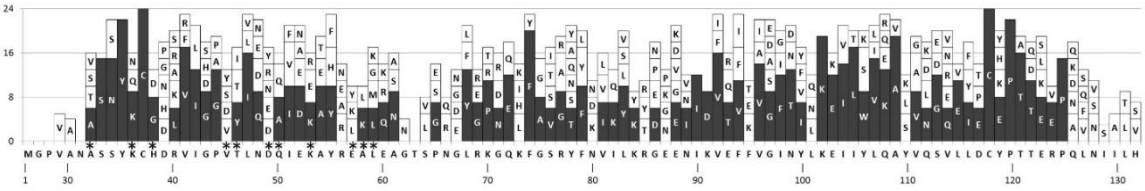


Figure 3. Conserved residues and motifs within the AVRPM3$^{A2/F2}$ family from wheat powdery mildew. Protein sequences of 24 family members without the signal peptide were aligned and compared to the AVRPM3$^{A2/F2}$ sequence. Residues present in at least three family members (white histograms) and the most frequent residues conserved in more than 6 of 24 family members (black histograms) are shown. Asterisks indicate residues under diversifying selection (probability of (ω>1) >0.95).

Including the two single mutations previously described (Bourras *et al.*, 2015), we generated 85 constructs individually mutating 66/106 residues (62.3%) with an average of 1.3 mutations per amino acid (Fig. 4a). We co-infiltrated each of these constructs with constructs expressing PM3A, PM3F, and the modified PM3F$^{L456P/Y458H}$ in *N. benthamiana* and assessed HR at 5 days post infection. Of the 85 single-residue altered constructs, 25 elicited HR to levels comparable to the wildtype ('neutral' alterations), while a majority of alterations (50/85, 59%) elicited no HR with any construct ('disruptive' alterations, Fig. 4a). For the alterations that were disruptive or neutral on PM3A and PM3F$^{L456P/Y458H}$ we did not observe HR when coexpressed with the natural PM3F allele. At four positions, amino acid changes were either neutral or disruptive, depending on the residue chosen for replacement (positions 52, 91, 106, and 119; Fig. 4a). Interestingly, we identified a single mutation (H132D) that abolished recognition by PM3A while maintaining the wildtype-level HR when co-expressed with PM3F$^{L456P/Y458H}$ (Fig. 4a,b). This is surprising since this is the final C-terminal residue of the AVRPM3$^{A2/F2}$ protein, and residues at the ends of proteins are not often implicated in protein-protein interactions or recognition specificity (Bos *et al,* 2010). Overall, we observed disruptive alterations across the AVRPM3$^{A2/F2}$ protein, while neutral alterations were concentrated in residues 60 to 75 (Fig. 4a).

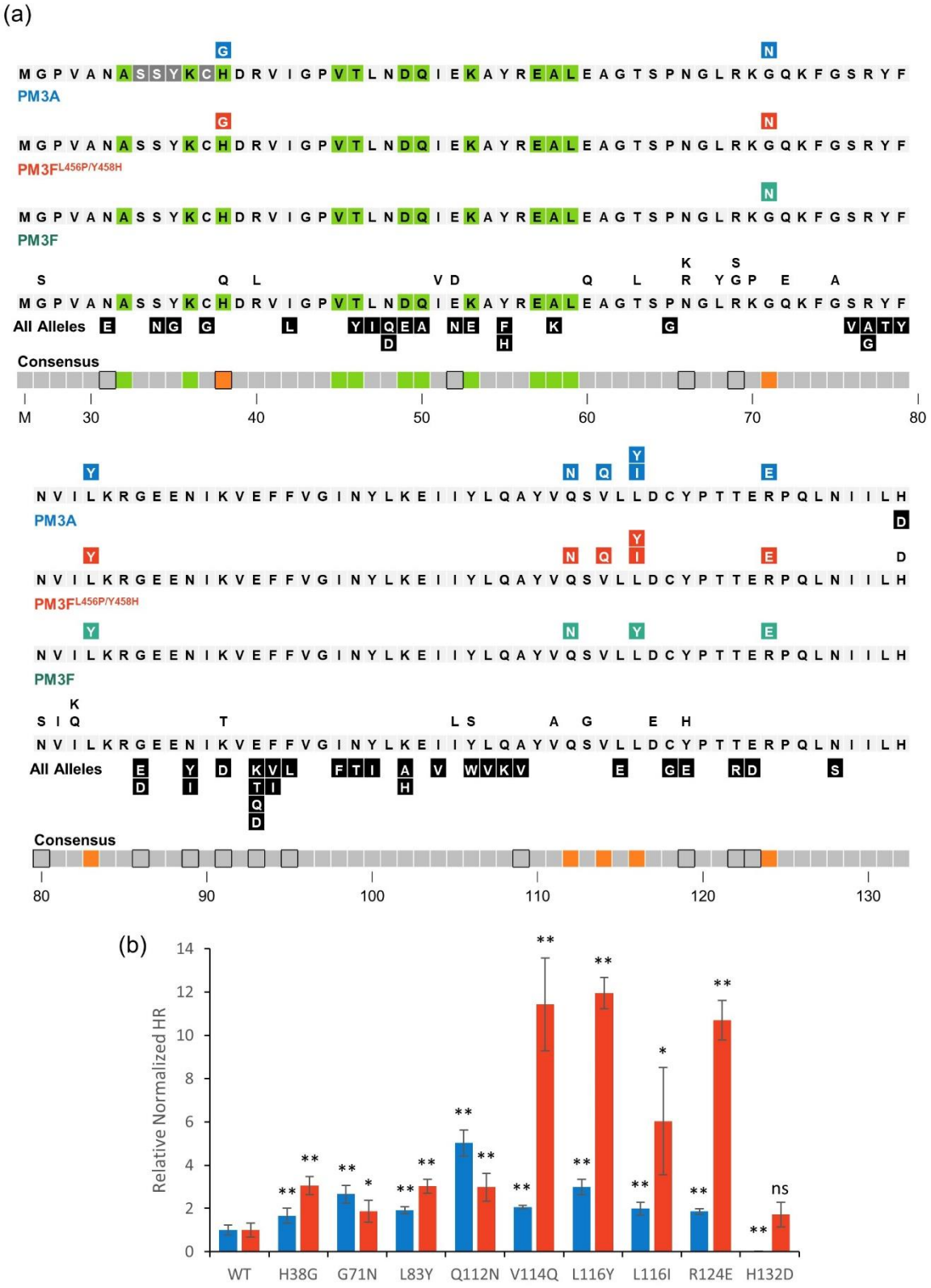


Figure 4. Single residue alterations in the wheat powdery mildew AVRPM3$^{A2/F2}$ avirulence protein tested by co-expression in *Nicotiana benthamiana* with the NLRs PM3A, PM3F, and PM3F$^{L456P/Y458H}$. Single residue changes derived from natural polymorphisms in the variants and sequence variation in the effector family span the complete AVRPM3$^{A2/F2}$ protein sequence without the signal peptide. These were tested by transient co-expression in *Nicotiana benthamiana* with constructs expressing PM3A, PM3F, and PM3F$^{L456P/Y458H}$. (a) The sequence of the AVRPM3$^{A2/F2}$ protein without signal peptide (replaced by methionine) is depicted, with the residues under *(continued on next page)*…

(continued)… diversifying selection in the family (green boxes) indicated. Results are shown on three lines separately for PM3A, PM3F$^{L456P/Y458H}$, and PM3F. Single residue alterations that do not affect recognition are not colored, while mutations that disrupt recognition (black boxes), enhance the hypersensitive response (HR) with PM3A (blue boxes), PM3F$^{L456P/Y458H}$ (red boxes) or gain a visible HR with PM3F (teal boxes) are indicated. A consensus diagram of the residues under positive selection in the family (green boxes), polymorphic residues in the variants (boxes outlined in black), and residues that result in enhanced HR with co-expressed with PM3A, PM3F, or PM3F$^{L456P/Y458H}$ compared to the wildtype (orange boxes) is given. (b) The subset of constructs selected for quantification based on their enhancement or altered patterns of HR when co-expressed with PM3A or PM3F$^{L456P/Y458H}$. Normalized quantification of the HR of mutants co-expressed with PM3A (blue histograms) or PM3F$^{L456P/Y458H}$ (red histograms) are shown relative to the wildtype. Asterisks indicate significant deviation (* and ** = $p<0.05$ and $p<0.01$, respectively) from the HR elicited by the wildtype construct individually for each leaf assay using a Two tailed Student T test. The error bars indicate the standard error of the mean (SEM).

Surprisingly, we identified eight mutations that enhanced the HR ('enhancing' alterations) with either PM3A, PM3F$^{L456P/Y458H}$, or both (H38G, G71N, L83Y, Q112N, V114Q, L116Y, L116I, and R124E; Fig. 4a). Due to the rapid HR responses, these reactions were quantified at 3 days post infection using fluorescence scanning (Praz *et al.*, 2017; Supporting Information Fig. S5), and we observed differences in the effects of each of the mutations on the cell death response by PM3A and PM3F$^{L456P/Y458H}$ (Fig. 4b). The alternations H38G, G71N and L83Y enhanced the cell death response by both PM3A and PM3F$^{L456P/Y458H}$ by 2- to 3- fold (Fig. 4b). For the Q112N construct, the increase in HR was stronger with PM3A (5-fold) than with PM3F$^{L456P/Y458H}$ (3-fold). The opposite was observed for the V114Q, L116Y, L116I and R124E mutations, where the increase in HR was stronger with PM3F$^{L456P/Y458H}$ than PM3A, and up to 12-fold higher than the wildtype for the V114Q, L116Y and R124E alterations (Fig. 4b). It is interesting to note that two positions identified as enhancing also gave different results depending on the residue alteration, where the change from Histidine to Glycine at position 38 enhanced recognition, but the change to Glutamine gave no discernable effect (Fig. 4a). Altering Leucine at position 116 to Tyrosine enhanced recognition by both PM3A and PM3F$^{L456P/Y458H}$ by 12-fold, while changing it to Isoleucine specifically enhanced PM3F$^{L456P/Y458H}$ recognition by 6-fold, and showed no effect on recognition by PM3A.

Previously it was shown that the wildtype AVRPM3$^{A2/F2}$-A variant induces only weak and occasional HR when co-expressed with the natural PM3F in *N. benthamiana* (Bourras *et al.*, 2015). Therefore we tested all single-residue altered constructs with PM3F to observe similarly enhanced recognition with the natural allele. We found five mutations that enhanced recognition with PM3A and PM3F$^{L456P/Y458H}$ and also elicited HR when co-expressed with PM3F to levels similar to the positive control (G71N, L83Y, Q112N, L116Y, and R124E; Fig. 5; Fig. 4a, teal boxes). A complete deletion of the region from Q112N to the L116Y mutation resulted in a disruption of recognition by PM3A, PM3F and PM3F$^{L456P/Y458H}$ (Fig. S6). Surprisingly, two mutations which gave a 12-fold increase in the response by PM3F$^{L456P/Y458H}$ (L116Y and R124E) also elicit HR with PM3F, however the Q112N alteration that is also recognized by natural PM3F was

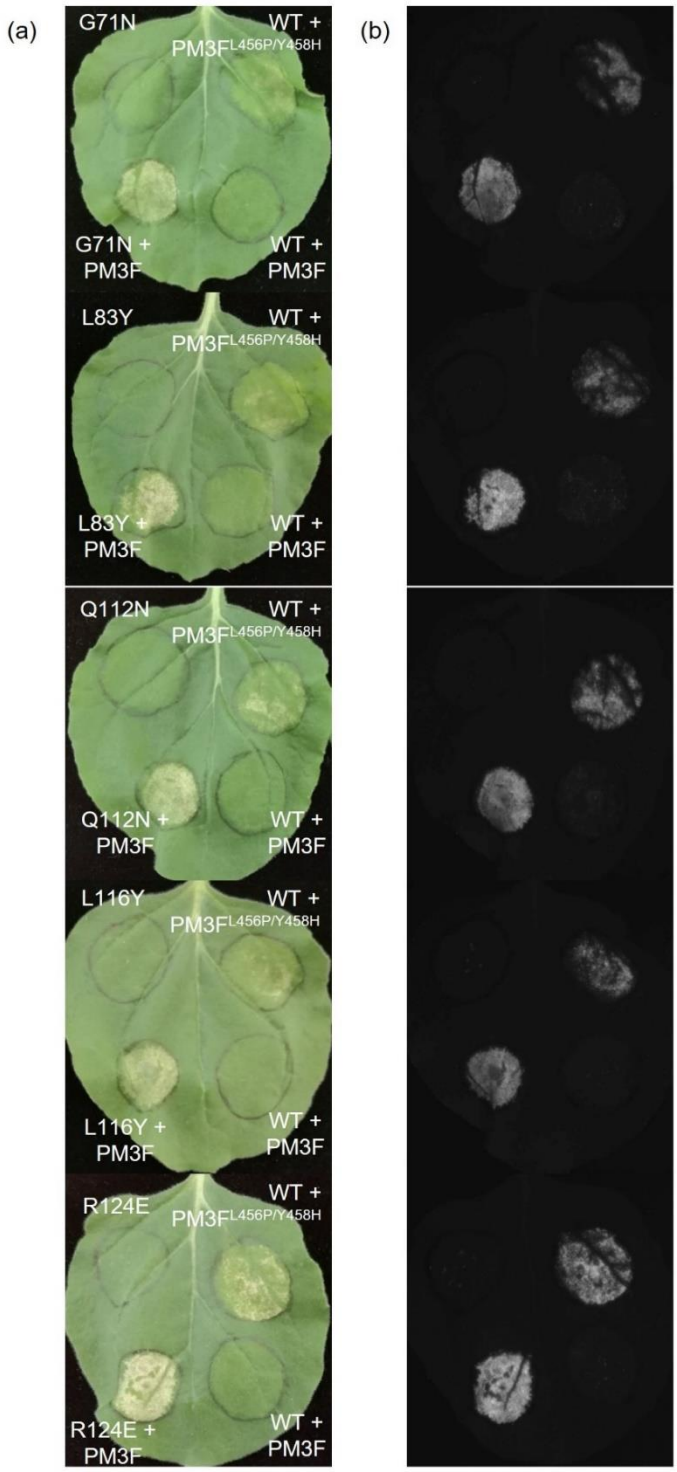


Figure 5. Single site mutations in the wheat powdery mildew AVRPM3$^{A2/F2}$ avirulence protein confer a gain of recognition when co-infiltrated with the natural PM3F allele. (a) Leaf images taken at 3 dpi demonstrate the strong hypersensitive response (HR) of the five single residue alterations that give enhanced recognition with the natural PM3F allele compared to wildtype (WT) AVRPM3$^{A2/F2}$. (b) Fluorescence imaging of the leaves visualizes the weak HR of the WT + PM3F control.

not so strongly recognized by the modified PM3F$^{L456P/Y458H}$, suggesting that the effect of these residues was specifically dependent on the L456P and Y458H modifications in the NBS domain of the PM3F$^{L456P/Y458H}$ protein.

We conclude that the cell death response following recognition of the AVRPM3$^{A2/F2}$ effector by PM3A, PM3F and the modified PM3F$^{L456P/Y458H}$ receptor is highly sensitive to single residue alterations in AVRPM3$^{A2/F2}$. We observed differences across the protein including a neutral region from residues 60-75, and a concentration of residue alterations resulting in an enhanced cell death response in the region from residue 112-116.

Comparing residues under selection pressure with the effects of single residue alterations on recognition by PM3A/F distinguishes putative AVRPM3$^{A2/F2}$ functional domains

To determine whether selection to evade recognition is a dominant driver of evolution among the *AvrPm3*$^{a2/f2}$ haplotypes, we analyzed the selection acting on the *AvrPm3*$^{a2/f2}$ gene and the *AvrPm3*$^{a2/f2}$ effector family. Among the 13 *AvrPm3*$^{a2/f2}$ haplotype sequences, there are a 6-bp insertion and 23 single nucleotide polymorphisms (SNPs), all but one of which are non-synonymous (Supporting Information Table S8). Excluding the *AvrPm3*$^{a2/f2}$*-G* haplotype, which is likely the result of a gene conversion event, the 278.4 nonsynonymous and 89.3 synonymous sites represent a significant excess of nonsynonymous substitutions compared to neutral expectation (P = 0.016). Tests for positive selection within the gene family using the Bayes empirical Bayes procedure (Yang *et al.*, 2005) to compare M8a and M8 models (see Materials and Methods) identified 11 positions under significant positive selection (Fig. 4a, green boxes). This region of the protein (residues 32-59) shares no overlap with the domains identified by single residue alterations as being distinctly neutral (residues 60-75) or enhancing recognition (residues 112-116). The positively selected residues also do not share substantial overlap with natural polymorphisms in *AvrPm3*$^{a2/f2}$ (Fig. 4a, black framed boxes). The single exception in both cases is position 38, where some haplotypes encode a neutral alteration to Glu (Q). The alteration to Gly (G) at this position, which is common within the effector family but was not found among natural AVRPM3$^{A2/F2}$ variants, caused 2-fold HR enhancement with PM3F$^{L456P/Y458H}$ and a 3-fold HR enhancement with PM3A (Fig. 4b).

Thus, we have identified a region under positive selection in the effector family (residues 32-59) and a protein domain where natural polymorphisms are concentrated (residues 86-95). Both regions are distinct from the two functionally identified domains, including the 'enhancing' region (residues 112-116), which might constitute an interaction domain. These data indicate that positive selection acting on the *AvrPm3*$^{a2/f2}$ effector family is distinct from the selection pressure driving diversity of *AvrPm3*$^{a2/f2}$ in natural isolates. This suggests that mutations to evade recognition by *Pm3a* and *Pm3f* would not necessarily interfere with effector function.

4.3 Discussion

Maintenance of a recognized effector at high frequencies in natural populations is possibly caused by the presence of a suppressor

Characterization of the natural diversity of fungal or oomycete avirulence genes has generally been limited to relatively small sample sizes (Allen *et al.*, 2008; Barrett *et al.*, 2009; Kanzaki *et al.*, 2012), or was primarily marker-based (Parlange *et al.*, 2009; Oliva *et al.*, 2015), a strategy which focuses exclusively on known polymorphisms and therefore cannot detect novel diversity. Our genotyping approach of sequencing complete genes enabled us to comprehensively identify the natural diversity of *AvrPm3$^{a2/f2}$*, globally. From this large collection we could observe that the recognized AVRPM3$^{A2/F2}$-A variant is maintained at high frequencies in all regions despite the presence of *Pm3a/f* wheat in cultivation. This contrasts with studies of *AvrLm4-7* diversity following prolonged exposure of a population of *L. maculans* to selection pressure from *Rlm7* resistance (Daverdin *et al.*, 2012). In the course of four years, the frequency of avirulent *AvrLm7* alleles sharply decreased and accordingly *Rlm7*-mediated resistance was all but defeated (Daverdin *et al.*, 2012). In our study, we observed the recognized AVRPM3$^{A2/F2}$-A variant at high frequencies, and many of these isolates also encoded the previously functionally validated variant of the suppressor-of-recognition, *SvrPm3$^{a1/f1}$-A* (Bourras *et al.*, 2015, Table 4, Supporting Information Fig. S3). The association we observe between active, expressed suppressor and virulence on *Pm3a* and *Pm3f* in these isolates fits with a model where the quantitative action of *SvrPm3$^{a1/f1}$* shelters *AvrPm3$^{a2/f2}$* from recognition by *Pm3a* and *Pm3f*.

However, our data also show that there are a considerable number of isolates that defy the three component *Avr-R-Svr* model (Table 4). First, we observed 24 isolates that do not encode the recognized AVRPM3$^{A2/F2}$-A variant but that nevertheless exhibit avirulence on *Pm3a*, *Pm3f*, or both (Table 2). More than half of these isolates (14) encode the AVRPM3$^{A2/F2}$-F variant and were originally collected on wild emmer wheat, hinting at the possible association between host diversity at the crop center of origin and pathogen diversity (Ben-David *et al.*, 2016). They have a genetic background closely related but distinct from isolates collected on domestic durum or bread wheat. There are likely additional factors in this genetic background that restrict virulence on the domesticated *Pm3a/f* wheat cultivars. The other ten exceptions might be explained by the presence of a second avirulence factor in the pathogen, recognized by *Pm3a*. This *AvrPm3^{a3}* was previously genetically identified in the cross of two European isolates (Bourras *et al.*, 2015), where segregating phenotypes in the progeny implied an additional recognized factor. It is likely that this factor is polymorphic in natural isolates, and leads to avirulent phenotypes in the absence of a recognized AVRPM3$^{A2/F2}$ variant.

In addition, we observed 16 isolates encoding AVRPM3$^{A2/F2}$-A and an SVRPM3$^{A1/F1}$ variant other than the functionally validated SVRPM3$^{A1/F1}$-A suppressor that exhibited some level of virulence on *Pm3a*, *Pm3f*, or both (Table 4). It is as yet unclear whether alternate variants of SVRPM3$^{A1/F1}$ such as the 'B' variant function as active suppressors. Follow up studies testing the suppressor activity of these variants and expanding expression studies to natural isolates from populations outside of Europe should determine the variability in suppressor activity of sequence-diverse variants and in natural populations globally.

Information from effector family sequence diversity can inform studies of recognition specificity

Previous mutagenesis approaches to investigate the basis of recognition specificity have focused primarily on limited polymorphisms in the natural diversity (Allen *et al.*, 2008; Parlange *et al.*, 2009; Kanzaki *et al.*, 2012) or random mutagenesis (Leonelli *et al.*, 2011), or have relied on information from already solved protein structures (Chou *et al.*, 2011; Blondeau *et al.*, 2015; Maqbool *et al.*, 2015). To reduce bias in selecting additional residue alterations outside the natural allelic diversity of the *AvrPm3$^{a2/f2}$* gene, we identified the most common residue at each position within the structurally related, sequence divergent family. Comparatively, Alanine screens have been widely used to test the role of specific residues in function and recognition (Joosten & de Wit, 1999; Whisson *et al.*, 2007). In theory alanine screening answers the question of the specific role of an amino acid side-chain, but the drastic alteration of the biochemical properties of residues is likely to also affect protein structure and stability. In contrast, by selecting common residues within the gene family toward which we altered single residues in the AVRPM3$^{A2/F2}$ protein, we reduced the chances of disturbing the overall structure of the protein. This is an important aspect to consider in a large-scale mutant screen where measuring protein stability for each tested construct is not feasible. This is particularly crucial for studies of the AVRPM3$^{A2/F2}$ protein, as attempts to tag the protein while maintaining functionality have thus far been unsuccessful. Ultimately, testing residues common within the effector family, which represents the pool of functional diversity available to the effector, not only reduces the probability of selecting residue alterations that may disrupt overall protein structure, but is also informative about the specificity of the receptor for the *AvrPm3$^{a2/f2}$* gene.

Utilizing an effector family-informed mutagenesis approach to study recognition specificity allowed us to identify single residue alterations that enhance the hypersensitive response upon recognition by the cognate R protein. To our knowledge, an enhanced hypersensitive response resulting from alterations in the AVR protein has not previously been described. None of the enhancing residue alterations we identified

exist as polymorphisms in the natural diversity of the *AvrPm3*$^{a2/f2}$ gene. Thus, a study limited to the natural sequence diversity, similar to *AvrLm4-7* (Blondeau *et al.*, 2015), *AvrL567* (Ravensdale *et al.*, 2012), and *ATR1* (Krasileva *et al.*, 2010; Chou *et al.*, 2011), would have found no enhanced recognition effects.

In addition, compared to random mutagenesis, our approach provided clues about which of the shared structural characteristics within the family are important for recognition. It enabled us to identify even conservative residue changes in the AVRPM3$^{A2/F2}$ protein that affect recognition specificity, and it also revealed unforeseen variation in the strength of the immune response by *Pm3a* and *Pm3f*. This was exemplified by the single biochemically conservative L116I alteration that resulted in a 10-fold increase in the HR elicited by the PM3F$^{L456P/Y458H}$ construct. It is possible that disruption and enhancement of the HR may be related to protein stability, however attempts to tag a functional AVRPM3$^{A2/F2}$-A have thus far been unsuccessful. Furthering our understanding of the role of specific residues in recognition will be a useful tool for the potential development of synthetic *R* genes with enhanced recognition abilities. Future studies of residues in the NLR that interact with the functional residues we have identified in this study might allow us to design a PM3 variant that recognizes the most common features or residues shared by the *AvrPm3*$^{a2/f2}$ family. Developing NLRs with such expanded structural recognition features would be a robust source of resistance in breeding.

Analyzing additional sequence diversity can distinguish effector-function domains from sites determining recognition specificity

By characterizing the natural diversity, testing single residue alterations for recognition, and estimating selection acting on the effector family, we have identified four distinct domains in the AVRPM3$^{A2/F2}$ protein (Fig. 6b): (i) residues under positive diversifying selection in the effector family (residues 32-59; Fig. 6b, green), that overlaps with the highly variable region and the YxC motif conserved in the effector family (Fig. 6a); (ii) the neutral domain (resides 60-75; Fig. 6b, dark grey) that also overlaps with the highly variable region; (iii) the domain that is most polymorphic in the *AvrPm3*$^{a2/f2}$ natural alleles (Fig. 6b, black frame); and (iv) the domain where residue alterations enhanced and/or altered the specificity of the HR elicited by PM3A/F (residues 112-116; Fig. 6b, orange). Assuming direct recognition of AVRPM3$^{A2/F2}$ by PM3A/F, the enhancing domain might constitute a putative interaction site. Traditionally, selection pressure from direct recognition by the NLR drives *Avr* diversification, often at the site of their protein-protein interactions (Ravensdale *et al.*, 2012; Maqbool *et al.*, 2015); however, we observed no naturally occurring polymorphisms at the putative interaction site. We did observe that most of the alterations in the naturally polymorphic domain disrupted recognition, but it is unclear

why this region shows a high level of diversification or whether this diversity is related to recognition, as most residue alterations across the protein are sufficient to disrupt recognition (Fig. 4a, black boxes). Since both of these domains are distinct from the domain under positive selection in the effector family, we conclude that the latter is likely involved in effector function, perhaps co-evolving with a target protein in the host.

In a model of indirect recognition, mutations that enhance effector function would increase the signal of a modified decoy or guardee that is detected by PM3A/F, leading to a stronger HR. Alternatively, if the effector function of AVRPM3$^{A2/F2}$ involves suppression of host defense responses and this activity is maintained in the *Nicotiana benthamiana* system, mutations in residues important for this virulence function might reduce the ability of the effector to suppress the hypersensitive response, but still allow for PM3A/F recognition. This would effectively enhance the visible HR in transient assays. Increased HR could also be the result of a change in the rate of effector protein turnover. Fine tuning the molecular interaction between an R protein that directly recognizes a cognate AVR or enhancing the indirect recognition of modifications to effector targets are both potential strategies for engineering enhanced resistance in this interaction.

Selection pressures other than recognition can also drive *Avr* allele diversification

Comparing the number of disruptive alterations based on natural diversity (10/16, 63%), from those designed using information from the effector family (50/74, 678%), we observed that polymorphisms identified in naturally occurring AVRPM3$^{A2/F2}$variants are equally as likely to disturb recognition as a selection of alterations from our screen. This contrasts strongly with the observation that the additive effects of multiple mutations are required to disrupt the direct recognition of *AvrL567* by *L5* and *L6* alleles (Ravensdale *et al.*, 2012). In the evolutionary context, one mutation that abolishes recognition is sufficient for eliminating host resistance as a source of selection pressure. Eight out of twelve natural AVRPM3$^{A2/F2}$ variants contain the G86E mutation, which is sufficient to disrupt recognition. Six of these variants have additional mutations that alone are sufficient to abolish recognition by PM3A/F (Table 1). The G86E polymorphism is therefore more likely to be ancestral than a gain-of-virulence mutation. In the remaining three variants (AVRPM3$^{A2/F2}$-C, -E, and –D) we find three single polymorphisms (F95L, E93D, and E93K, respectively) representing putative gain-of-virulence mutations. The remaining six natural polymorphisms most likely either existed prior to *Pm3a/f* recognition, are the product of random genetic drift, or are the result of selection pressure from other sources, such as effector function. This hypothesis is supported by the absence of significant positive selection

acting at positions within the putative interaction domain (residues 112-116). We detected, however, significant positive selection at residues in the effector family, which should result from selection for effector function and not recognition. Novel effector functions have been described in *Cladosporium fulvum*, where both *Avr4* and *CfEcp6* effectors bind chitin through their Lysin (LysM) domains, but only *Avr4* was demonstrated to have the additional function of protecting fungal hyphae from hydrolysis by plant enzymes (van den Burg *et al.*, 2006; de Jonge *et al.*, 2010). All of the mildew isolates tested here encode the *AvrPm3*$^{a2/f2}$ gene, suggesting loss of the gene might result in fitness costs. Additionally, the region experiencing positive diversifying selection in the effector family, characteristic of shared effector function, is not the domain experiencing diversification in the *AvrPm3*$^{a2/f2}$ alleles. We conclude that neither recognition nor shared effector family function but rather an important novel function that increases pathogen fitness is potentially the main driver of diversity within the *AvrPm3*$^{a2/f2}$ gene.

4.4 Materials and Methods

Fungal collection, propagation, and virulence tests

Blumeria graminis isolates were maintained on detached leaves of the appropriate susceptible cereal cultivar: 'Kanzler' (*Triticum aestivum*), 'Inbar' (*Triticum durum*) or 'Matador' (*Triticosecale*); on benzimidazole agar as described by Parlange *et al.*. (2011). The worldwide collection contained 272 isolates, 55 isolates collected in 13 states in the United States, 101 collected in 12 provinces in China (Zeng *et al.*, 2014), 61 collected in 4 eco-geographic regions of Israel (Ben-David *et al.*, 2016), 51 collected primarily in Switzerland as well as in five other European countries (our collection), 2 in Japan, and 2 in Australia. All isolates from Europe, Japan, Australia, and a subset of isolates from the US, China, and Israel were maintained in house. Spore samples for DNA extraction and sequencing by PCR were obtained primarily for US and Israeli isolates, while complete genomes for a subset of Chinese isolates were used for *in silico* extraction (Praz *et al.*, 2017). European, Japanese and Australian isolates were collected in the field and single spore isolated twice as described by Brown and Wolfe (1990). Virulence tests on *Pm3a* (cv. Asosan/8*Chancellor) and *Pm3f* (cv. Michigan Amber/8*Chancellor) of isolates in house were performed as described by Brunner *et al.* (2010), and scored as described in Bourras *et al.* (2015). External Israeli, US, and Chinese isolates were collected, maintained, and phenotyped as described in Ben-David *et al.* (2016), and Zeng *et al.* (2014), respectively.

DNA/RNA isolation and construction of plasmid vectors

High molecular weight DNA for PCR amplification was extracted as previously described by Bourras *et al.* (2015). RNA samples were extracted from infected detached 'Chancellor' leaves using the Qiagen miRNeasy Mini Kit (Qiagen) according to the

manufacturer. Full-length cDNA was prepared using the Superscript III RT kit (Invitrogen) according to the manufacturer. Molecular cloning into Gateway-compatible entry vector was performed using the pENTR/D-TOPO Cloning Kit (Invitrogen) according to the manufacturer. Gene synthesis including Gateway compatible cloning sites was performed by Gen9 (https://www.gen9bio.com/), resulting in Gateway compatible cloning vector (pG9m-2) containing the synthesized gene. The synthesized gene was then cloned directly into the pIPKb004 expression vector (Himmelbach *et al.*, 2007) used for transient expression in *N. benthamiana*. Site-directed mutagenesis was performed by PCR amplification using overlapping primers containing desired mutations on the Gateway compatible pENTR/D-TOPO vector containing *AvrPm3*$^{a2/f2}$. Recombination of the mutant PCR product into the binary vector pIPKb004 was performed as previously described by Stirnweis *et al.* (2014). Primers used for gene amplification and site directed mutagenesis are listed in Supporting Information Table S1. Constructs are listed in Supporting Information Table S2. All *AvrPM3*$^{a2/f2}$ and *SvrPm3*$^{a1/f1}$ genomic sequences used for this study are available at the GenBank database under the accession numbers MG739404 - MG739429.

Genetic analyses and tests for selection

The Templeton, Crandall and Sing (TCS) network was visualized using the PopArt software (Leigh & Bryant, 2015). Maximum Likelihood computations and estimation of the average nonsynonymous/synonymous (d_N/d_S) rate ratio for *AvrPm3*$^{a2/f2}$ haplotypes was conducted using HyPhy software package (Pond *et al.*, 2005) in the MEGA7 program (Kumar *et al.*, 2016). All multiple alignments were performed with Muscle 3.8.31 (Edgar, 2004). The protein alignments were back translated to nucleotide alignments using TranslatorX v1.1 (Abascal *et al.*, 2010). To test for positive selection, we estimated the likelihood of the maximum likelihood tree under the M8a and M8 model (Yang *et al.*, 2000) with Paml 4.8 (Yang, 2007) and using the Bayes Empirical Bayes method (Yang *et al.*, 2005).

Quantitative real-time PCR experiments

RT-qPCR experiments were performed as described in Praz *et al.* (2017) with the modification of a 20 second extension time. Three independent biological replicates were sampled at 2 days after inoculation. Glyceraldehyde 3-phosphate dehydrogenase (*Gapdh*) was used as an internal control as described in Bourras *et al.* (2015). Gene expression was normalized to that of *Gapdh*. RT-qPCR primers for reference and target genes were previously described (Bourras *et al.*, 2015).

Transient protein expression assays in Nicotiana benthamiana

Transient expression by agroinfiltration in *Nicotiana benthamiana* was conducted according to the protocols of Ma *et al.*. (2012) and Dugdale *et al.*. (2014), and with the modifications described by Bourras *et al.* (2015). To test for *AvrPm3-Pm3* interaction,

Agrobacterium expressing the *AvrPm3* constructs or the *Pm* resistance allele were mixed in a 4:1 ratio of *Avr:R*. At least three replications were performed and the hypersensitive response (HR) was visualized as described by Praz *et al.* (2017). HR intensity was calculated using the mean gray value estimated by ImageJ (Schneider *et al.*, 2012) and normalizing against the non-infiltrated background. Normalized values were compared between HR elicited by mutant $AvrPm3^{a2/f2}$ constructs and wildtype $AvrPm3^{a2/f2}$ constructs, and statistical significance was estimated using the Students T test.

Acknowledgements

This work was supported by the Swiss National Science Foundation (grant 310030-163260). We would like to acknowledge Prof. Dr Kentaro Yoshida (Kobe University, Kobe, Japan), Prof. Dr Tomohiro Ban (Yokohama City University, Yokohama, Japan), and Dr Simon Ellwood (Curtin University, Perth, Australia) for collecting and shipping mildew isolates. We would also like to acknowledge Simon Flückiger and Michael Schläfli for their technical support throughout the revisions.

Author Contributions

K.E.M., S.B., and B.K. designed the study. K.E.M., L.L., and L.K. performed the experiments. K.E.M., F.M., and C.R.P. performed analyses. K.E.M., C.R.P, M.C.M, R.B.D., K.C., A.D., and E.M. collected data. K.E.M., S.B., and B.K. wrote and edited the manuscript. R.B.D., C.C., M.X., F.Z., S.G. and D.Y. contributed data.

CHAPTER 5

Population Specific Evolution of Diverse Strategies to Overcome *Pm3a/f* recognition of the *AvrPm3$^{a2/f2}$* Avirulence Factor

Kaitlin Elyse McNally[1], Fabrizio Menardo[1], Patrick Christoph Brunner[2], Marion Claudia Müller[1], Christoph Stritt[1], Roi Ben-David[3], Christina Cowger[4], Dazhao Yu[5,6,7], and Beat Keller[1]

[1]Institute of Plant and Microbial Biology, University of Zürich, Zollikerstrasse 107, 8008 Zürich, Switzerland
[2]Department of Environmental Systems Science, ETH Zürich, Universitätstrasse 2, 8092 Zürich, Switzerland;
[3]Institute of Plant Science, ARO-Volcani Center, 50250 Bet Dagan, Israel;
[4]United States Department of Agriculture-Agricultural Research Service (USDA-ARS), Department of Plant Pathology, North Carolina State University, Raleigh 27695, USA;
[5]Institute of Plant Protection and Soil Science, Hubei Academy of Agricultural Sciences, 430064 Wuhan, China;
[6]Ministry of Agriculture Key Laboratory of Integrated Pest Management in Crops in Central China, 430064 Wuhan, China;
[7]College of Life Science, Wuhan University, 430072 Wuhan, China.

Unpublished data

Supplemental Data in Appendix III

Summary

The quantitative action of SVRPM3$^{A1/F1}$ in the *AvrPm3-Pm3-SvrPm3* interaction was originally described in a cross of two European isolates, but was later found in multiple isolates collected in Europe. It remains unclear whether this three-component model applies to other geographic populations. To answer this question we compared between four geographic populations *AvrPm3$^{a2/f2}$* and *SvrPm3$^{a1/f1}$* gene expression, signatures of selection pressure acting on *AvrPm3$^{a2/f2}$*, and the diversity and functional properties of natural polymorphisms in *AvrPm3$^{a2/f2}$* alleles. We found no evidence to suggest suppression is acting outside of the European population. We observed signatures of a recent bottleneck in the European population, while in the Israeli population there was evidence of balancing selection maintaining a uniquely diverse set of *AvrPm3$^{a2/f2}$* alleles. We identified single polymorphisms in AVRPM3$^{A2/F2}$ variants unique to the United States which may represent gain-of-virulence mutations not observable in the other populations. Taken together, we conclude that differences in population histories, stability, and structure, as well as dispersal patterns, have led to population-specific evolution of the *AvrPm3-Pm3* interaction.

5.1 Introduction

Studies of disease dynamics and local adaptation in natural ecosystems have been informative for the development of theories of coevolution of plants and their pathogens (Burdon and Thrall, 2009; Bousset and Chèvre, 2013). Host-pathogen coevolution depends upon several factors, including both pathogen and host demographics, life histories, spatial and temporal variation, and environment (Burdon and Thrall 2009; Barrett et al, 2008; 2009; Laine et al, 2011). Due to this complexity, an important approach in the study of coevolution is analysis at the metapopulation level, rather than evaluating coevolution in single spatial or temporal populations (Blanquart et al, 2013). Metapopulation studies have described the evolutionary tradeoff between transmission and infectivity of pathogens (Bull, 1994; Thrall and Burdon, 2003), the influence of gene flow, as well as the role of genetic diversity (Thrall and Burdon, 2000; Laine, 2005) and overall resistance of host populations (Laine et al, 2014). Thus far studies have been inherently host-focused, and generalized to the average level of resistance and susceptibility of a host population (Laine et al, 2014; Barrett et al, 2015). However, this generalized approach is not informative for host-pathogen coevolution in agricultural systems, where monocultures of strongly resistant cultivars predominate (Bousset and Chèvre, 2013). Therefore, because both limited host genetic diversity and heightened host resistance are known to substantially increase pathogen evolution (Thrall and Burdon, 2003; Niemi et al, 2006), accelerated host-pathogen coevolution in agricultural systems is likely to fundamentally differ from coevolution in natural systems.

There are relatively few metapopulation studies of host-pathogen coevolution in agrosystems (Bousset and Chèvre, 2013). This is surprising given the concepts and technologies that have arisen out of links drawn between plant genetics, epidemiology, population genetics and agronomy, such as durable disease resistance selection (Johnson, 1984), the concept of pathogen evolutionary potential (McDonald and Linde, 2002; Stuckenbrock and McDonald, 2008), and integrated crop protection (Krupinsky et al, 2002). Recent insights from the study of the recognition specificity of single interactions between host resistance (*R*) genes and pathogen avirulence (*Avr*) genes have led to the development of several concepts applying to these specific and dynamic interactions. First, repeated exposure to single *R* gene containing cultivars in an agricultural setting is sufficient to drive immense genetic diversification of cognate *Avr* alleles in the pathogen, often detected as positive selection (Daverdin et al, 2012). Second, the mechanistic variation of recognition specificity influences the type of *Avr* allelic diversity generated. For example, direct recognition of AVR proteins by R proteins has been shown to drive sequence diversification in both AVR and R proteins (Kanzaki et al, 2012). By contrast, indirect recognition through AVR effector function is thought to drive gene loss or inactivation, resulting in presence/absence polymorphisms in the pathogen population (Van der Hoorn et al, 2002). However, loss of important AVR effector proteins can sometimes result in a significant decrease of pathogen fitness (Huang et al, 2006; 2010), leading to an evolutionary tradeoff for the pathogen. Because of these contrasting evolutionary forces acting on *Avr* loci, it is as yet unclear what determines the observed evolutionary mechanism leading to gain-of-virulence in the pathogen, whether it be *Avr* gene loss or sequence diversification.

The diversity of gain-of-virulence mechanisms has been well described in effectors from the hemibiotrophic ascomycete fungi, *Magnaporthe oryzae* (rice blast) and *Leptosphaeria maculans* (stem canker). Virulent isolates of these pathogens on several resistance genes in rice (*Oryae sativa*) and rapeseed (*Brassica napus*), respectively, have been most often linked to *Avr* gene deletions and only more rarely to residue polymorphisms (Huang et al, 2014; Yoshida et al, 2016; Rouxel and Balesdent, 2017). Despite the general use of coevolutionary concepts to explain allelic diversity and to generate hypotheses, thus far studies of *Avr-R* coevolution in these pathosystems have been extremely limited. A four year field trial to study the evolution of virulence at the *AvrLm4-7* locus of a population of *L. maculans* in response to *Rlm7* resistance resulted in a dramatic increase in the frequency of virulent isolates over time (Daverdin et al, 2012). Analysis of virulent *AvrLm4-7* alleles revealed a higher proportion of point mutations leading to inactive *AvrLm4-7* alleles initially, but the most common gain-of-virulence allele observed at the conclusion of the trial was complete gene deletion (Daverdin et al, 2012). These results suggest characteristics of virulent *Avr* alleles may change over time, however, this study did not take into account other environmental factors that might influence *Avr* evolution in spatially distinct populations.

AvrPm3$^{a2/f2}$ is a secreted effector protein of the obligate biotrophic fungal pathogen, *Blumeria graminis* f. sp. *tritici* (*B.g. tritici*), the causal agent of wheat powdery mildew. *AvrPm3*$^{a2/f2}$ is specifically recognized by the *a* and *f* alleles of the *Pm3* allelic series, which confer overlapping race-specific resistance to *B.g. tritici* isolates (Bourras et al, 2015 [Chapter 3]). It has been observed that *Pm3a/f* recognition of *AvrPm3*$^{a2/f2}$ is quantitatively suppressed by a suppressor-of-recognition, *SvrPm3*$^{a1/f1}$ (Bourras et al, 2015 [Chapter 3]), suggesting that the outcome of *Pm3* resistance is determined according to an *Avr-R-Svr* model of interaction (Bourras et al, 2016). Additionally, higher expression of *SvrPm3*$^{a1/f1}$ relative to *AvrPm3*$^{a2/f2}$ expression in five European *B.g. tritici* isolates was shown to be positively associated with virulence on *Pm3a*, and to a higher degree, *Pm3f* (Bourras et al, 2015 [Chapter 3]). In fact, most of the evidence of a quantitatively acting suppressor is observed in European isolates, as compared to isolate populations from the US, China, and Israel (Chapter 3), leading to the question of whether the *Avr-R-Svr* model is generally valid.

In this study, we analyzed *SvrPm3*$^{a1/f1}$ and *AvrPm3*$^{a2/f2}$ expression in isolates from the Israeli population exhibiting different phenotypes on cultivars expressing *Pm3a* and *Pm3f* to observe whether there is evidence of quantitative suppressor action. We compared estimates of selection acting on *AvrPm3*$^{a2/f2}$ between US, European, Israeli, and Chinese mildew populations and compared the results with observations of *AvrPm3*$^{a2/f2}$ allelic diversity within each population. From our results and observations we conclude that different selection pressures have led to unique *AvrPm3*$^{a2/f2}$ allelic diversification within distinct populations, and hypothesize that the evolution of a suppressor-of-recognition was unique to the European population, and that sequence divergence was the main gain-of-virulence mechanism in isolates in the US population. Based on these conclusions, we propose that pathogen evolutionary history and demographics have a strong influence on *Avr-R* coevolution, possibly rivaling the role of the mechanism of recognition specificity in determining the evolution of pathogen virulence.

5.2 Results

Expression analysis of Israeli isolates reveals evidence of unique modifiers

Previously we analyzed expression levels of *AvrPm3*$^{a2/f2}$ and *SvrPm3*$^{a1/f1}$ in four European isolates exhibiting either avirulent or virulent phenotypes on the *Pm3a* allele, revealing an association between higher *SvrPm3*$^{a1/f1}$ expression and virulence on *Pm3a* (Chapter 3). In the phenotypes of the isolates harboring both an active *AvrPm3*$^{a2/f-2}$*-A* and *SvrPm3*$^{a1/f1}$*-A* in the worldwide collection, we observed no evidence of suppressor activity in isolates from the US, but phenotypic evidence of the quantitative action of the suppressor is observable in isolates from China as well as Israel (Table 1). To test

whether suppression acts in isolates from Israel, we examined expression of *AvrPm3*$^{a2/f2}$ and *SvrPm3*$^{a1/f1}$ in three isolates from Israel, one avirulent/avirulent on *Pm3a/Pm3f*, one virulent/avirulent, and one virulent/virulent and compared the expression patterns to the four isolates from Europe (Figure 1). We observed no association between expression of *SvrPm3*$^{a1/f1}$ in the Israeli isolates and their phenotypes on *Pm3a/f*. Together with the lack of evidence of the activity of an *Svr* in the US population, this suggests factors besides *SvrPm3*$^{a1/f1}$ might be contributing toward the *AvrPm3*$^{a2/f2}$*-Pm3a/f* interaction and that other mechanisms to evade recognition might be utilized in other populations.

Table 1. Isolates from the Israeli, Chinese, and US populations carrying both the active *AvrPm3*$^{a2/f2}$ and active *SvrPm3*$^{a1/f1}$ genes.

	Origin	AVRPM3$^{A2/F2}$	SVRPM3$^{A1/F1}$	*Pm3a* wheat	*Pm3f* wheat
OKS(14)-B-3-2	US	A	A	A	V
OKH-B-3-3	US	A	A	A	
Isr204	Israel	A	A	A	A
Isr4	Israel	A	A	A	V
Isr68	Israel	A	A	A	V
Isr70	Israel	A	A	A	V
Isr7	Israel	A	A	V	V
Isr94	Israel	A	A	V	V
Isr101	Israel	A	A	V	V
Isr215	Israel	A	A	V	V
45-10	China	A	A	A	V
46-25	China	A	A	A	V
46-31	China	A	A	A	V
47-3	China	A	A	A	V
48-18	China	A	A	A	V
SD-5	China	A	A	A	V
1-19	China	A	A	A	
13-50	China	A	A	A	
30-1	China	A	A	A	
50-2	China	A	A	A	
9-2	China	A	A	A	
GZ-10	China	A	A	A	
GZ-6	China	A	A	A	
36-70	China	A	A	V	V
NZ-1	China	A	A	V	V
11-99	China	A	A	V	
9-10	China	A	A	V	

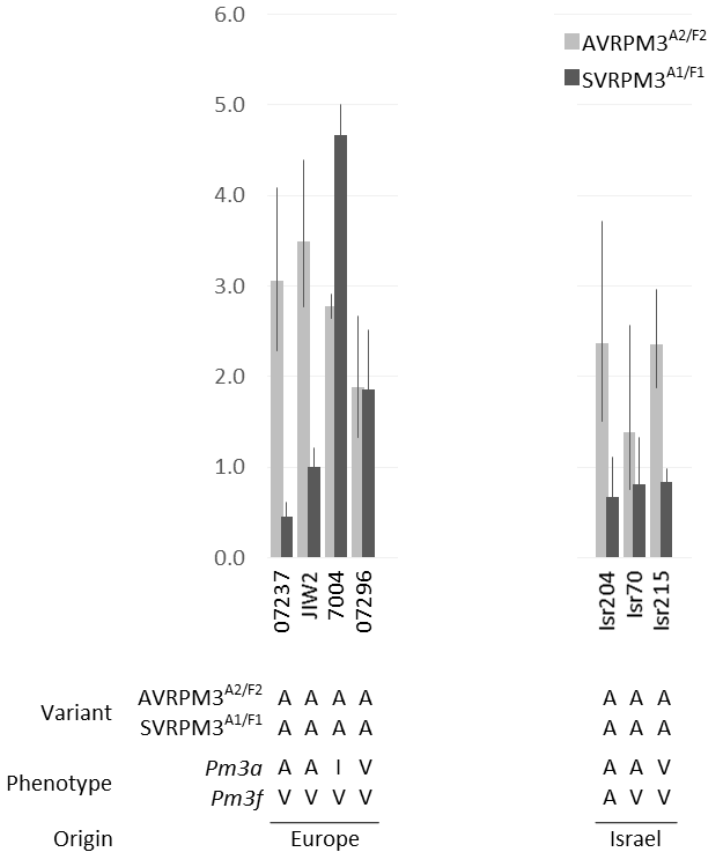


Figure 1. Expression analysis of isolates encoding the AVRPM3$^{A2/F2}$-A and SVRPM3$^{A1/F1}$-A variant combination and which give different phenotypes on *Pm3a*. The mean normalized expression of *AvrPm3*$^{a2/f2}$ (light grey histograms) and *SvrPm3*$^{a1/f1}$ (dark grey histograms) at 2 days post infection. Leaf segments from the susceptible recurrent parent line '*Chancellor*' were infected with *B.g. tritici* isolates 07237, JIW2, 7004 and 07296 from Europe and Isr204, Isr70 and Isr215 isolates from Israel that contain the same *Avr-Svr* genotype and are virulent on *Pm3f*, but show different virulence (avirulent, A; intermediate, I; virulent, V) on *Pm3a* or *Pm3f*.

Different selection pressures act in distinct isolate populations

Natural sequence diversity among AVR variants is thought to result from positive selection to evade recognition by R proteins in a manner reflecting the mechanism of interaction with cognate R proteins (Van der Hoorn et al, 2002; Dodds et al, 2006). To examine whether other mechanisms to evade recognition by *Pm3a/f* might be evident in other populations we compared the amount of natural diversity of the AVRPM3$^{A2/F2}$ variants within each population. The diversity of haplotypes in Israel far exceeded the diversity of haplotypes from other regions (Figure 2A). We found 7 variants in the Israeli population, 5 in China, 4 in Europe and 3 in the US (Figure 2A). Not only does the US population contain the least number of variants, but two of these variants are

unique to the US and contain only single polymorphisms from the recognized AVRPM3$^{A2/F2}$-A variant, also present in the population (Figure 2B). Not only does this population contain fewer variants, but it is clear from the Templeton, Crandall and Sting (TCS) haplotype network that variants in this population are less diverse, either representing recent divergence or a selective sweep.

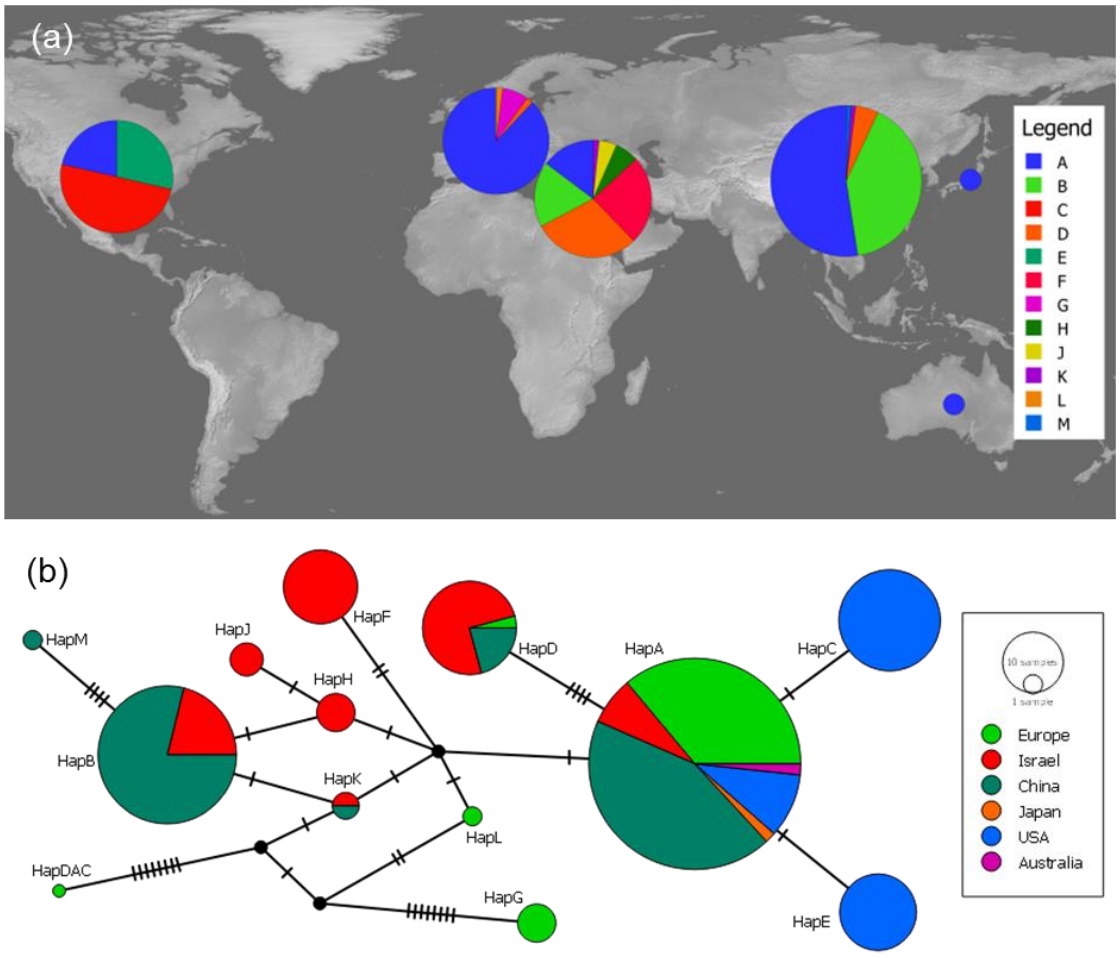


Figure 2. Haplotype network and geographic distribution of the AVRPM3$^{A2/F2}$ variants. (a) Geographic distribution of the 12 AVRPM3$^{A2/F2}$ variants (A-M) from the US, Europe, Israel, China, Japan, and Australia. Area of the circles correspond to the number of isolates from that region. (b) Haplotype network of 12 *AvrPm3*$^{a2/f2}$ haplotypes (HapA-HapM), and the *B.g. dactylis* ortholog (HapDAC) colored by geographic origin. Area of the circles correspond to the frequency of each haplotype. Small black circles represent missing haplotypes in the network. Notches on the connections indicate number of nucleotide polymorphisms between sequences.

Spatially distinct populations can experience different selection pressures acting on *Avr* loci, which would contribute toward unique natural alleleic diversity between populations. For example, both 'arms-race' and 'trench-warfare' dynamics are thought to lead to different patterns of natural diversity in *Avr* alleles (Van der Hoorn et al, 2002; Thrall et al, 2012). To estimate selection pressure acting within the populations we performed the Tajima's D test of neutrality (*D*) for *AvrPm3*$^{a2/f2}$ (Table 2, See Appendix III: Supplemental table 1). Estimates above 2 or below -2 are considered

significant signals of selection (Tajima, 1989). Interestingly we observed different signatures of selection between populations, most significantly evidence of a selective sweep in Europe (D = -2.04) and balancing selection in Israel (D = 2.48). Population structure is known to interfere with estimates of Tajima's D.

Table 2: Overview of haplotype diversity by origin and the Tajima's D test of neutrality (D) for *AvrPm3*$^{a2/f2}$.

	Isolates	A	A-T	B	C	D	E	F[a]	G[b]	H	J	K	L	M	*D*
USA	56	12	-	-	28	-	16	-	-	-	-	-	-	-	1.68
Europe	50	41	3	-	-	1	-	-	(4)	-	-	-	1	-	-2.04
Israel	61	9	-	11	-	18	-	(15)	-	4	3	1	-	-	2.48
China	101	53	-	41	-	5	-	-	-	-	-	1	-	1	-0.22
Globally	272	119	3	52	28	24	16	(15)	(4)	4	3	2	1	1	-0.43

[a]Excluded from analysis because of population structure
[b]Excluded from analysis because of insertion mutation

To check whether the US isolates demonstrate population structure at the level of the *AvrPm3*$^{a2/f2}$ gene, we mapped the origins of isolates carrying the three variants (Figure 3). Isolates carrying the "A" variant are distinctly separated from isolates carrying either the "C" or "E" variant. Therefore it is likely that strong population structure skews the estimate of Tajima's D for the US population.

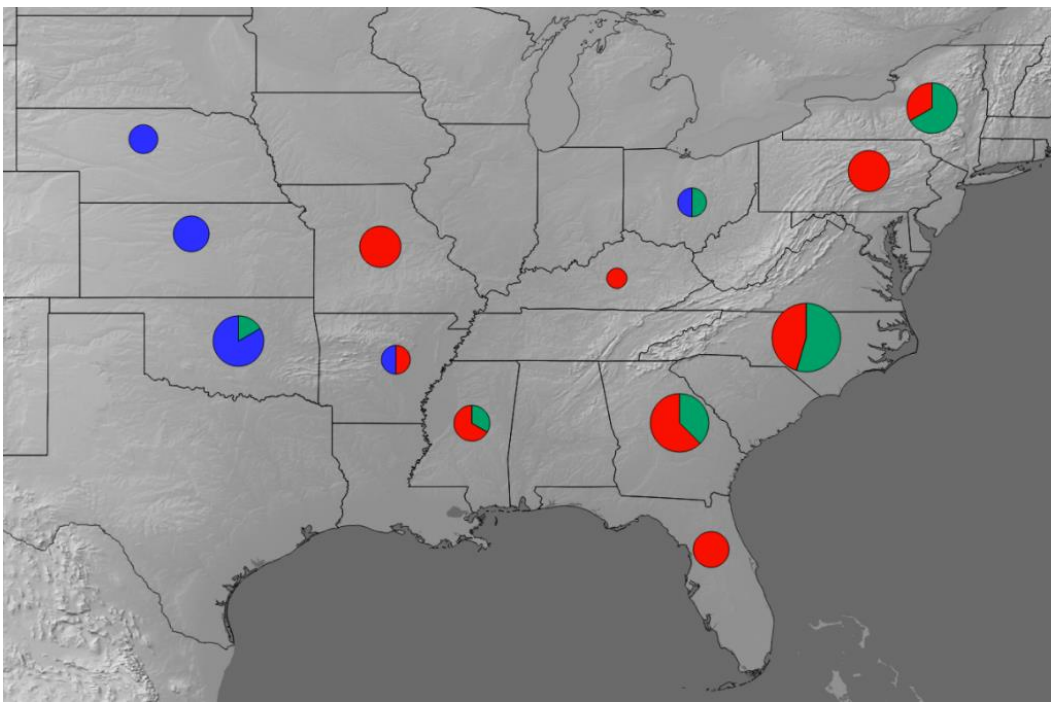


Figure 3: Distribution of isolates in the US carrying the AVRPM3$^{A2/F2}$ -A, -C, and –E variants.
Geographic distribution of the AVRPM3$^{A2/F2}$ -A (blue), -C (red), and –E (green) variants across the US. Area of the circles correspond to the number of isolates from that state.

To summarize, Israel harbors the most natural diversity of the populations in our collection and is experiencing balancing selection between these variants. This is in contrast to European isolates, which appear to be experiencing a selective sweep or recent evolutionary bottleneck that has resulted in a high frequency of a single AVRPM3$^{A2/F2}$ variant, the recognized "A" variant. We conclude from these results that selection is acting differently in distinct populations, leading to diverse patterns of natural diversity in *AvrPm3$^{a2/f2}$*.

Natural diversity of the AVRPM3$^{A2/F2}$ variants suggests alternate mechanisms to evade recognition

Distinct patterns of natural diversity between mildew populations might reflect alternative mechanisms of *Avr-R* coevolution and gain-of-virulence mutations. Looking specifically to the natural diversity of variants within populations (Table 3), we observe that the most polymorphic variants can be found in Israel, China, and Europe, while the two least polymorphic (C and E) are exclusive to the US. Orthologs from related species can provide useful information to estimate the age of specific polymorphisms in the natural diversity. Comparing the diversity in the AVRPM3$^{A2/F2}$ variants against the close ortholog from *B.g. dactylis* (AVRPM3$^{A2/F2}$-Dactylis; Menardo et al, 2017; Figure 4; Table 3), we can associate three of the most common polymorphisms to the ortholog. This suggests that these residues might have been shared in the ancestral progenitor. Interestingly, one of these polymorphisms, G86E, has been previously shown in transient assays to disrupt *Pm3a/f* recognition (Chapter 3). Assuming this polymorphism is ancestral, then it is unlikely that other polymorphisms in these variants arose from selection pressure to evade recognition. There are three exceptions, E93K, E93D and F95L, that could potentially have arisen from previously recognized haplotypes in response to selection pressure from *Pm3a/f* recognition. These three mutations might therefore represent gain-of-virulence mutations, two of which are specific to variants from the US population. We conclude from these observations that in agreement with previous findings (Chapter 3), polymorphisms in the natural diversity from the European, Israeli, and Chinese populations are not the result of selection pressure to evade recognition, whereas single polymorphisms in the two unique variants in the US population likely represent gain-of-virulence mutations driven by selection pressure from *Pm3a/f* resistance.

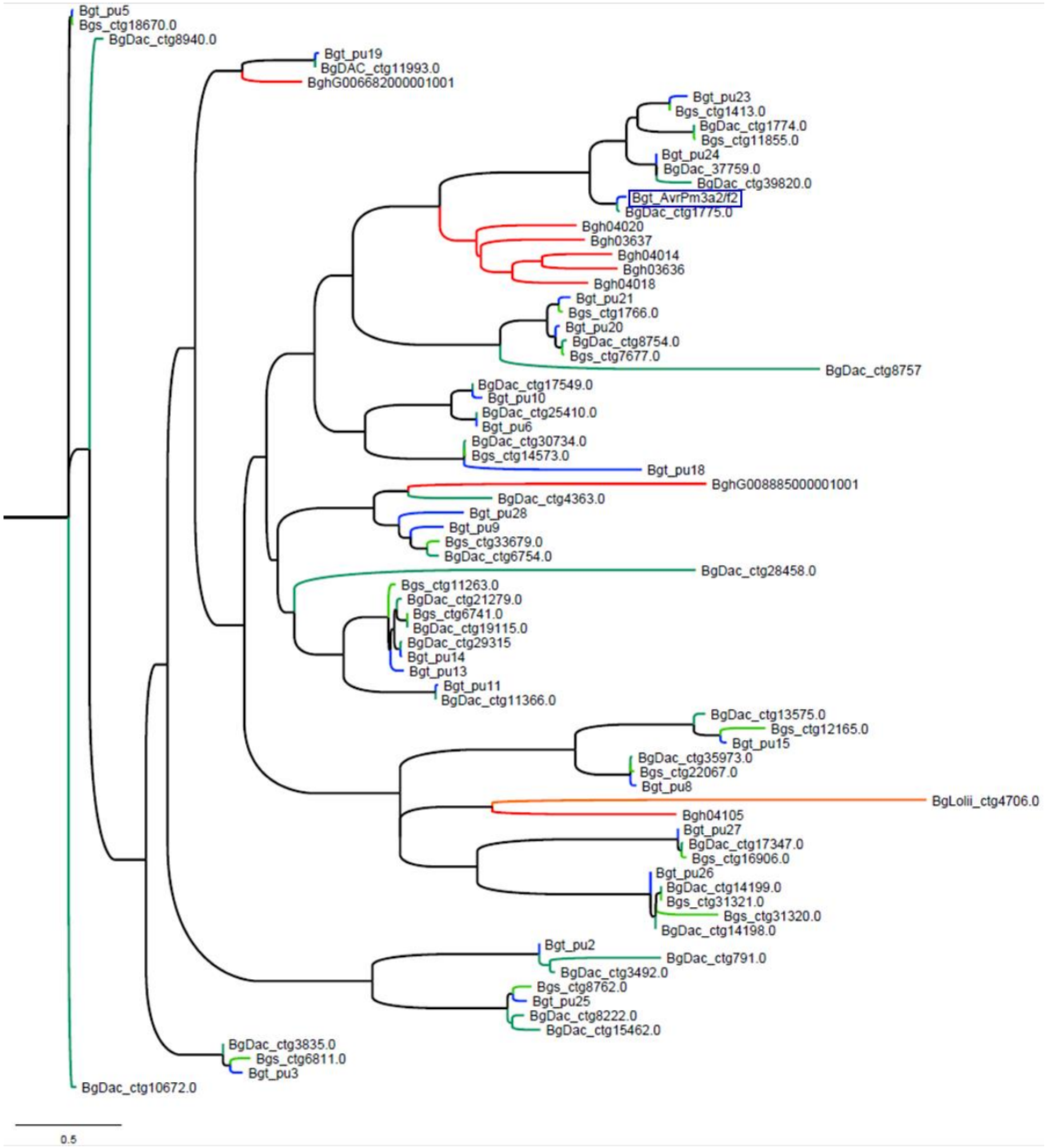


Figure 4. Phylogenetic tree of the *AvrPm3*$^{a2/f2}$ family in *B.g. tritici* and its homologous families in *B.g. hordei, B.g. secalis, B.g. dactylis* and *B.g. lolii*. Maximum likelihood tree of the *AvrPm3*$^{a2/f2}$ family in *B.g. tritici* (blue branches) and its homologous families in *B.g. hordei* (red branches), *B.g. secalis* (green branches), *B.g. dactylis* (teal branches), and *B.g. lolii* (orange branch). The homologous families in *B.g. dactylis* and *B.g. secalis* are more closely related to the *B.g. tritici* family than *B.g. hordei* or *B.g. lolii*. A direct homolog of *AvrPm3*$^{a2/f2}$ (blue box) was found only in *B.g. dactylis*.

Table 3. Disruptive mutations in the amino acid sequences of the AVRPM3$^{A2/F2}$ variants.

	No.	7	12	21	24	25	26	27	31	38	52	66	69	80	86	89	91	93	95	109	112	114	119	122	123
AVRPM3$^{A2/F2}$-A[a]	**122**	**V**	**I**	**A**	**S**	**G**	**-**	**-**	**N**	**H**	**E**	**N**	**R**	**N**	**G**	**N**	**K**	**E**	**F**	**A**	**Q**	**V**	**Y**	**T**	**E**
AVRPM3$^{A2/F2}$-C	28	.	.	.	.	.	-	-	.	.	.	.	.	.	.	.	.	.	L	.	.	.	.	.	.
AVRPM3$^{A2/F2}$-E	16	.	.	.	.	.	-	-	.	.	.	.	.	.	.	.	.	D	.	.	.	.	.	.	.
AVRPM3$^{A2/F2}$-F[b]	15	.	.	.	.	.	-	-	.	.	.	K	S	.	E	.	.	.	.	.	.	.	.	.	.
AVRPM3$^{A2/F2}$-D	24	.	.	.	.	.	-	-	.	.	.	K	.	S	.	.	T	K	.	.	.	.	.	.	.
AVRPM3$^{A2/F2}$-B	52	.	.	.	.	.	-	-	.	Q	.	.	.	.	E	.	.	.	.	.	.	.	.	.	D
AVRPM3$^{A2/F2}$-G	4	.	.	V	N	S	P	V	E	Q	.	K	.	.	E	Y	.	.	.	V	.	.	H	.	.
AVRPM3$^{A2/F2}$-M	1	.	.	.	.	.	-	-	E	Q	N	.	.	.	E	.	.	.	.	.	.	.	.	.	D
AVRPM3$^{A2/F2}$-K	2	.	.	.	.	.	-	-	.	Q	.	.	.	.	E	.	.	.	.	.	.	.	.	.	.
AVRPM3$^{A2/F2}$-H	4	.	.	.	.	.	-	-	.	.	.	.	.	.	E	.	.	.	.	.	.	.	.	.	D
AVRPM3$^{A2/F2}$-J	3	.	.	.	.	.	-	-	.	.	.	.	.	.	E	.	.	.	.	.	.	.	.	R	D
AVRPM3$^{A2/F2}$-L[a]	1	.	.	.	.	.	-	-	.	.	.	.	.	.	E	.	.	.	.	V	.	.	.	.	.
AVRPM3$^{A2/F2}$-Dactylis	1	M	V	.	.	.	-	-	.	Q	.	K	.	.	E	D	.	.	.	.	L	Q	.	.	.
Total	273																								

Polymorphisms compared to the avirulent variant are depicted using the one-letter code for amino acids. Residue numbering includes the signal peptide. Number of isolates encoding each variant is given (No.). Residues that individually disrupt recognition by PM3A and PM3F$^{L456P/Y458H}$ are indicated with black boxes. Untested residues are in grey boxes. Dashes indicate gaps and dots indicate identical residues.

[a] Bourras et al, 2015

[b] Variant found in isolates collected on *T. dicoccoides* (wild emmer)

5.3 Discussion

Our study of the patterns of selection and natural diversity on the metapopulation level has revealed distinct differences within each population, which we will here describe and attempt to interpret. The European population is where the study of the *AvrPm3-Pm3* interaction began, and where we first provided evidence for the quantitative action of the suppressor of recognition, *SvrPm3$^{a1/f1}$* (Bourras et al, 2015 [Chapter 3]). Interestingly, the low *AvrPm3$^{a2/f2}$* haplotype diversity and higher frequency of rare alleles indicate that the European population might have recently experienced a selective sweep or recent bottleneck that brought the AVRPM3$^{A2/F2}$-A variant to a high frequency. This is striking given that the AVRPM3$^{A2/F2}$-A variant is still strongly associated with avirulence in this population (72%), despite the action of the suppressor. This is again evidence that this variant might confer a strong fitness benefit to the pathogen that out-competes selection pressure from recognition by *Pm3a/f*.

In the US population, where we would expect intensive *R* gene deployment to increase selection for evading recognition in pathogen populations, we found evidence of deviation from random mating that might indicate an interaction between *Avr* and *Svr*, as well as evidence of balancing selection acting on *AvrPm3$^{a2/f2}$* haplotypes. However,

both of these findings are most likely explained by strong population structure, which was recently described by Cowger and colleagues (2016). In their study, they show that isolates in the Southern Plains region, where *Pm3a* is still an effective source of resistance, form a distinct genetic cluster from isolates to the East, where *Pm3a* has been defeated since the 1990s (Niewoehner and Leath, 1998). This division into two genetically distinct populations is almost perfectly associated with the *AvrPm3$^{a2/f2}$* haplotype sequence (Figure 3). Additionally, the two unrecognized variants AVRPM3$^{A2/F2}$-C and AVRPM3$^{A2/F2}$-E are unique to the US, and contain only one residue alteration each, both of which are sufficient to avoid recognition. We conclude that these two variants likely arose in the eastern US population as a response to selection pressure by intensive *Pm3a* usage. This suggests that although the active SVRPM3$^{A1/F1}$ is widely present in this population, it might not have been sufficiently expressed to suppress recognition by *Pm3a*, "forcing" the pathogen to mutate AVRPM3$^{A2/F2}$ to evade recognition. This is supported by the fact that the US is the only population where we find a 100% association between presence of the AVRPM3$^{A2/F2}$-A variant and avirulence on *Pm3a*.

The deviation from random mating that we observed in the Chinese population is also likely to be due to population structure, as observations of virulence frequencies for a variety of *Pm* genes showed differences between provinces, ranging from 19.9 to 52.8% for *Pm3a* (Zeng et al, 2013). Estimates of selection pressure were not significant for either *AvrPm3$^{a2/f2}$* or *SvrPm3$^{a1/f1}$* in Chinese isolates, but varying selection pressures between isolated populations would obscure an average estimate. A more detailed analysis of population structure in isolates from China will be necessary to draw any conclusions about the selection acting in this region.

The Israeli population is unique and worth closer investigation for several reasons. Israel is part of the geographic region where cereals originated. Early studies of *B.g. tritici* isolates from this region suggested that coevolution during the extreme diversification of grasses led to a correspondingly high genetic diversity within cereal powdery mildews (Eshed and Wahl, 1970). Today, domestic and wild grasses grow in close proximity, and in a recent study of differentiation among isolates from Israel Ben-David and colleagues (2016) found genetic differentiation between wild and domestic *B.g. tritici* groups. The natural diversity of *AvrPm3$^{a2/f2}$* and *SvrPm3$^{a1/f1}$* fit this model, because we find unique haplotypes among isolates collected from wild cereals (*AvrPm3$^{a2/f2}$-F* and *SvrPm3$^{a1/f1}$-E*). We also find a strong association between the other three most common *AvrPm3$^{a2/f2}$* haplotypes, *-A, -B,* and *-D*, and isolates collected on domestic cereals. Ben-David and colleagues (2016) also reported that the domestic *B.g. tritici* isolates in the Israeli population show no significant genetic population structure. We find strong evidence of balancing selection maintaining the frequencies of the three *AvrPm3$^{a2/f2}$* hexaploid haplotypes in the domestic isolates.

It is unlikely that all virulent AVRPM3$^{A2/F2}$ variants are derived from the recognized AVRPM3$^{A2/F2}$-A by selection pressure to evade recognition, as many of the shared polymorphisms among these variants are common with the ortholog AVRPM3$^{A2/F2}$-Dactylis, from the closely related *B.g. dactylis forma specialis*. This suggests that these shared residues are more ancient than the residues in the AVRPM3$^{A2/F2}$-A variant. We propose that the AVRPM3$^{A2/F2}$-A, -B, and -D variants originated in Israel following wheat domestication, where they appear to be maintained under balancing selection. We also propose that the AVRPM3$^{A2/F2}$-A variant evolved toward conferring an enhanced fitness on the pathogen, which would explain its prevalence in the sampled populations of Europe and China.

In the natural populations, we observed on one hand, evidence of balancing selection in the Israeli population between several AVRPM3$^{A2/F2}$ variants, although only one of them is recognized by any *Pm3* allele. This balancing selection might have some other origin, such as an effector target that is diverse among the many *Triticum* species growing in close vicinity in Israel.

Conclusions and Next Steps

This is the first time multiple large populations from around the world have been assessed for natural diversity at high enough frequency to allow for characterizing their specific mechanism of recognition specificity. This has revealed four population specific patterns of diversity and potential interaction models. In their review of the *Leptosphaeria maculans* pathosystem, Rouxel and Balesdent (2017) argue that *Avr* evolution is not pathogen-specific, but gene-specific. We would like to add that we find evidence that *Avr* evolution might not only be gene-specific, but population specific, and determined by population histories, stability, and structure, as well as dispersal patterns (Barrett et al, 2008; Burdon and Thrall, 2009). As we observe, individual populations may evolve distinct mechanisms to alter or evade recognition specificity and this can be influenced by factors outside of the AVR-R interaction.

CHAPTER 6

General Discussion

6.1 Do the *AvrPm3$^{a2/f2}$* effector family genes encode intein-like protein structures?

AVRPM3$^{A2/F2}$ shares no structural homology to fungal ribonucleases, unlike the well characterized RALPH effectors, and the recently reported AVR$_{a1}$, AVR$_{a13}$ and AVRPM2 from *Blumeria graminis* (Spanu, 2014; Lu et al, 2016; Praz et al, 2017). AVRPM3$^{A2/F2}$ does share structural homology to the sequence divergent effector family of which it is one of 24 members. We performed a preliminary structural modeling analysis of the *AvrPm3$^{a2/f2}$* effector family (Appendix IV). This revealed significant estimates of a common secondary structure shared by the family, and indicated that despite significant divergence in protein sequences, the family shares an overall secondary structure. The secondary structure prediction is also shared by the most frequent template used by the protein threading algorithm of RaptorX, the intein domain of an intein homing endonuclease from *Thermococcus kodakarensis*. However, since none of the structural predictions passed significance thresholds, the most convincing aspect of this result is that the template is shared between more than half of the family members (Appendix IV: Table A1). Inteins are internal protein elements capable of self-splicing post-translationally to form the mature protein product, and they are found in all three domains of life (Novikova et al, 2014). Inteins are closely related to the fundamental Hedgehog (Hh) signal transduction pathway, which highlights their ancient origin and the diversity of their enzymatic functions (Bürglin, 2008). Inteins without the optional endonuclease domain (mini-inteins) range in size from 130-200 amino acids and are well known for having very high levels of sequence divergence outside of several highly specific motifs crucial for the enzymatic reaction that confers their characteristic autoproteolytic function (Mills et al, 2014). One necessary motif is two Cys residues, one at the N-terminal and one at the C-terminal of intein domains (Eryilmaz et al, 2014), a feature that is 100% conserved among the *AvrPm3$^{a2/f2}$* effector family. A search for other intein motifs in the effector family resulted in the identification of a potential derivative of one of the most important intein motifs, TxxH, which is located near the middle of smaller inteins (Eryilmaz et al, 2014). We found a similarly located TxxK motif in more than half of the family members, however several other requisite motifs could not be identified. Mutations in the TxxK motif disrupt recognition by PM3A/F, but were inconclusive regarding their enzymatic function. This is similar to RALPH effectors, which are structurally related to ribonucleases but do not encode a ribonuclease active site (Spanu, 2017). It should be possible to screen the TxxK mutants

for altered effector function once an assay to test effector virulence function has been established.

6.2 Reasoning behind a split-intein hypothesis

The *AvrPm3*$^{a2/f2}$ family ranges in size from 90-111 amino acids after signal peptide cleavage. The full intein domain of the intein homing endonuclease from *Thermococcus kodakarensis* is in two parts, from residues 1-120 and 501-537 (PDB: 2cw7; Matsumura et al, 2006). Therefore, the family does not cover the entire intein domain of the intein homing endonuclease structure. Additionally, the TxxK motif in the family resembling the TxxH intein motif is shifted towards the C-terminus, relative to where it is found in mini-inteins (Eryilmaz et al, 2014). We therefore explored the literature for variant intein structures which might more closely resemble our model. One possibility is the N-intein domain of a 'split'-intein. Split-inteins are divided into the N- and C-intein domains that are encoded separately, but then bind to one another *in vivo* to form a complete intein and carry out protein splicing *in trans* (Shah et al, 2013). N-intein domains have been described to range in size from 99-123 residues (Shah and Muir, 2014). Additionally, the region of the N-intein that specifically binds the C-intein partner has been described (Shah et al, 2013). This region in the AVRPM3$^{A2/F2}$ structural model (Figure 1A) has significant overlap with the KxILxVKA conserved motif (Appendix IV: Appendix Figure A.1) and residues affecting recognition specificity (Figure 1B). To date inteins have not been described in multi-cellular organisms, but are well documented in fungi, including several ascomycetes (Bokor et al, 2012; Novikova et al, 2014). Split-inteins occur rarely (Shah and Muir, 2014), but their splice site and molecular splice mechanisms can be highly divergent (Bachmann and Mootz, 2015). Most intriguingly, they have been proposed to be agents of protein domain shuffling following multiple cross-reacting splicing events in *trans* (Aranko et al, 2013; Figure 1C). Theoretically this process would allow the formation of a variety of protein products from *trans* splicing events between multiple intein-like effectors (Pietrokovski, 2001; Figure 1D). One hypothesis is that the *AvrPm3*$^{a2/f2}$ effector family might have been derived from inteins and are now secreted effectors that undergo *trans* splicing in the host cell to bypass detection by host immune receptors. The final products of these splicing events would then be the functional effectors.

The most conservative hypothesis of the intein model is that the current effector function no longer resembles an intein function. Inteins are believed to be the remnants of the earliest and most basic enzymes (Novikova et al, 2014). As with the RALPH effectors, the *AvrPm3*$^{a2/f2}$ effector family might share an overall structure with inteins, but might have developed a novel, unrelated function that is not immediately evident given our current knowledge and experimental capabilities. Solving the protein structure and studying how mutations effect virulence function will be important areas of further study to understand this novel effector family.

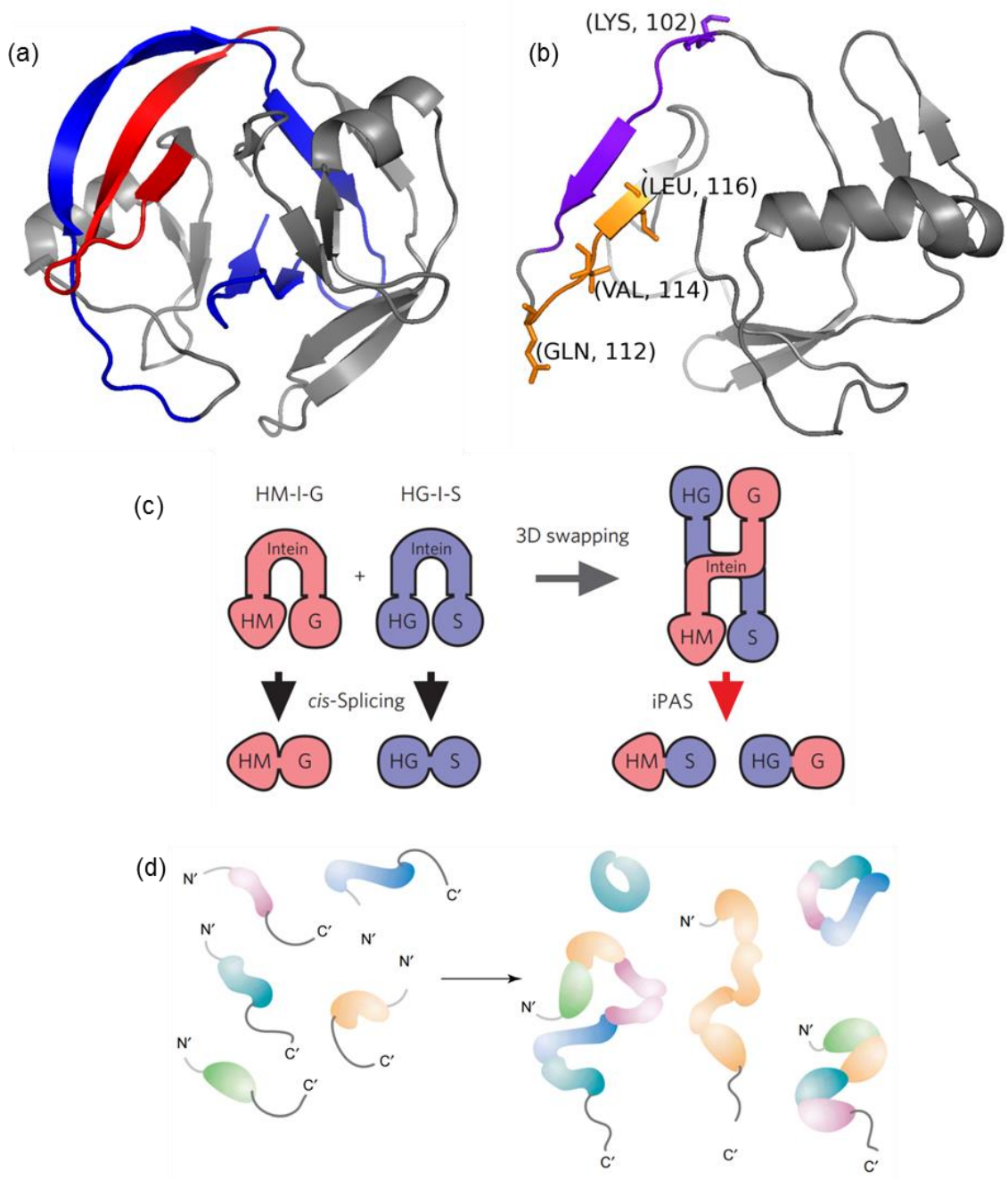


Figure 1. Comparison of functionally important residues in the *AvrPm3*$^{a2/f2}$ family and the residues involved in intermolecular interactions between the N- and C-intein of the DnaE split-intein complex from *Nostoc punctiforme*. (a) Crystal structure of a split intein (Oeemig et al, 2009), divided into the N-intein (grey and red) and C-intein (blue). The region of the N-intein responsible for binding the C-intein consists of two β-strands and a connecting loop (red). (b) The AVRPM3$^{A2/F2}$ protein sequence modeled onto the predicted intein structure of the closest relative, *Pu_23*, with the region corresponding to the highly conserved KxILxVKA motif (purple) and the residues that enhanced the HR response or altered recognition specificity (orange, labels) comprise two β-strands and a connecting loop. (c) Schematic drawing of two *cis* splicing intein precursors (HM-I-G and HG-I-S) producing two alternatively ligated products (HM-S and HG-G) taken from Aranko et al (2013). (d) Schematic drawing of combinatorial *trans* splicing intein precursors. Colored ovals represent protein segments with flanking N- and C-terminal intein parts depicted as black lines. Taken from Pietrokovski (2001).

6.3 Support for a direct interaction hypothesis

There are seven functionally distinct alleles (*a-g*) of the *Pm3* resistance gene that have been cloned from *T. aestivum*, and an additional 10 (*k-t*) were cloned from the tetraploid *T. durum* and hexaploid wheat land races (Yahiaoui et al, 2004; Srichumpa et al, 2005; Bhullar et al, 2009; 2010). The *Pm3a, b, c,* and *f* alleles are distinct from other alleles because of their overlapping race-specificities (*Pm3b/c* and *Pm3a/f*), and because they contain polymorphic sequence blocks rather than SNPs present in other alleles (Brunner et al, 2010). The region most polymorphic between all *Pm3* alleles from *T. aestivum* is LRR 25 to 28, immediately following the unusually large and flexible LRR 24, and it was therefore predicted to be the site of direct interaction with the cognate AVR protein (Sela et al, 2014). Previously we demonstrated that the *Pm3a* and *Pm3f* alleles share specificity for $AvrPm3^{a2/f2}$, which is also not recognized by the closely related *Pmb/c* alleles, or the more distant *Pm3d, e,* or *g* alleles (Bourras et al, 2015 [Chapter 3]). If we look for associations in the polymorphisms between alleles, two regions of the LRR are identical in *Pm3a/f* but differ from *Pm3b/c*, and those are LRRs 2-4 and LRRs 26 and 27. Together this supports the model proposed by Sela *et al* (2014) of direct interaction with the AVR in LRRs 26-28 determining specificity. Additionally, according to the putative intein structural model, residues conferring important specificity of recognition by *Pm3a/f* are highly surface exposed, which is consistent with a model of direct interaction between AVR and R proteins.

6.4 Support for an indirect interaction hypothesis

However, the evolution of the *Pm3a, b, c,* and *f* alleles does not fit the usual pattern of direct interaction. Direct interaction is typically characterized by an 'arms-race' dynamic where multiple polymorphisms in both *R* and *Avr* genes is observed. This is perfectly exemplified by the interaction of alleles of *AVR-Pik* from *Magnaporthe oryzae* and alleles of the *Pik* resistance gene from rice (*Oryzae sativa*). The physical interaction of AVR and R proteins drives selection in both genes and a step-wise evolution of recognition and evasion by single residue alterations is clearly observable (Kanzaki et al, 2012). In contrast, the *Pm3* alleles *a* and *f* both recognize only one $AVRPM3^{A2/F2}$ variant, and the *Pm3b* and *Pm3c* alleles recognize a sequence-unrelated *Avr*, as none of the other 23 family members are recognized by any other *Pm3* allele. That leaves open the possibility of an indirect recognition model. The limited *Pm3* allelic diversity might reflect evolution towards monitoring a guardee that is shared by the *Pm3a/f* allele pairs, while another is shared by *Pm3b/c*. Examples of *R* genes detecting modifications to their guardee by effector activities have been described (Rooney et al, 2005). Specificity, in this case, might be achieved through the modifications to the binding affinity of the AVR for the guardee, and the stability of this complex formation. Alternatively, in the case of an enzymatic function, partial modification or slowed enzymatic alteration of

the guardee due to residue variation in the AVR might result in intermediate protein products that delay or alter the strength of NLR signaling. We identified putative enzymatic motifs using the intein structural model, however alterations of these putative motifs to either enhance or eliminate intein enzymatic activity both resulted in a loss of recognition by PM3A/F, suggesting that if enzymatic activity does occur at these residues, then it is not involved in an indirect recognition response by PM3A/F.

While the polymorphisms of the *Pm3a, b, c,* and *f* alleles do not match the pattern expected for direct interaction, the polymorphisms in *AvrPm3$^{a2/f2}$* are also ambiguous in regard to the mechanism of this interaction. Indirect recognition is usually associated with presence/absence polymorphisms of the *Avr* which distinguish between virulent and avirulence races, while direct recognition is associated with high genetic diversity in the *Avr* protein (Van der Hoorn et al, 2002; Jones and Dangl, 2006; Xiao et al, 2008). We observed no isolates lacking a putatively functional *AvrPm3$^{a2/f2}$* haplotype, however we also observed limited natural sequence diversity within haplotypes and hypothesize that most of the diversity is not the result of selection pressure to evade recognition. It is likely that the involvement of other pathogen-encoded factors such as the suppressor of recognition (*SvrPm3$^{a1/f1}$*) in this interaction mask the typical evolutionary clues as to the mechanism of *AvrPm3$^{a2/f2}$* recognition by *Pm3a/f*.

6.5 Interpreting population-specific evolution

Wheat powdery mildew has followed close behind the cultivation of wheat ever since domesticated bread wheat left the Fertile Crescent. The most recent introduction of wheat cultivation occurred in North America following European colonization. Therefore we assume that the US powdery mildew population is the youngest population in our collection, while Israel has likely the oldest population. Life history is known to play an important role in the evolutionary potential of host-pathogen interactions (Barrett et al, 2008). The recent introduction of a small population of mildew isolates constituting limited genetic diversity combined with the intense evolutionary pressure of industrial agriculture typical in the US likely enhanced selection pressure on *AvrPm3$^{a2/f2}$* to evade recognition by *Pm3a/f*. This has been observed in recent years by the defeat of *Pm3a* and *Pm3f* containing wheat cultivars (Cowger et al, 2016). Indeed, we observe single nucleotide polymorphisms in each of the two AVRPM3$^{A2/F2}$ variants (C and E) unique to the US population that disrupt recognition by both *Pm3a* and *Pm3f*. The distinct geographic distribution of recognized and unrecognized variants in the US population is strong evidence to suggest that these SNPs are part of a selective sweep that defeated *Pma/f* in the Eastern US (Cowger et al, 2016). Looking exclusively at the population of US isolates, we might conclude that *Pm3a/f* and *AvrPm3$^{a2/f2}$* co-evolve in an arms-race dynamic driven by direct interaction at the two polymorphic sites in the AVRPM3$^{A2/F2}$-C and -E variants.

However, comparing the natural diversity of *AvrPm3$^{a2/f2}$* in the US against the diversity in Israel suggests that such a conclusion would be an over-simplification. The relatively more ancient and genetically diverse mildew population from Israel is not experiencing intensive agricultural practices, but is exposed to a more dynamic climatic and genetically diverse hosts (Ben-David et al, 2016). In the natural diversity of this population we do not observe causative mutations that would suggest direct interaction with *Pm3a/f*, nor do we find evidence of selective sweeps. Instead, we find high frequencies of multiple alleles indicative of balancing selection. Balancing selection is characterized as rotating alleles that are maintained because of conflicting forces, which in the case of a pathogen *Avr* is most likely fitness and avirulence. The high frequency of 4 AVRPM3$^{A2/F2}$ variants in Israel cannot exclusively be the result of frequency-dependent selection from the interaction with the resistance gene since only one of the variants is recognized by any *Pm3* allele. Additionally, we found that the AVRPM3$^{A2/F2}$-F variant occurs only in isolates collected from wild emmer, which are less virulent on hexaploid wheat (Ben-David et al, 2016). There is evidence that wild emmer isolates are a source of diversity for domestic mildew isolates (Ben-David et al, 2016), however we do not observe AVRPM3$^{A2/F2}$-F variants in domestic isolate genetic backgrounds. Taken together, these are further evidence that balancing selection acting on *AvrPm3$^{a2/f2}$* is indicative of interactions with an effector target in the host, which is likely to be more polymorphic in Israel than in other regions, rather than the interaction with *Pm3a/f*.

It is possible then, that the high frequency of the recognized AVRPM3$^{A2/F2}$-A variant is also related to the genetic diversity of cultivated wheat varieties in other growing regions. The effect of the active suppressor of recognition, *SvrPm3$^{a1/f1}$-A*, which when highly expressed relative to *AvrPm3$^{a2/f2}$-A*, effectively suppresses recognition by *Pm3a/f*, is most observable in isolates from the European population. In spite of the expected fitness benefits of expressing the active AVRPM3$^{A2/F2}$-A variant while effectively evading recognition, we do not observe a high frequency of isolates (11/23, 47.8%) which carry both variants but evade recognition by *Pm3a* and *Pm3f*. Perhaps the selection pressure of *Pm3a/f* use in Europe is not strong enough to induce a selective sweep, or perhaps we are observing the intermediate stage of such a sweep. Another possibility is that there is a fitness cost associated with high expression of the suppressor. Further studies of the changing frequencies of these alleles and assays to assess their contribution to fitness will help to answer this question, which will be important for designing effective resistance strategies against virulent mildew races.

6.6 Implications of complexity for future breeding strategies

As advances in our knowledge of recognition specificity bring us closer to being able to design resistance genes, the question becomes how to effectively use this technology

and what would it target? Relaxing (or enhancing) specificity of *Pm3a/f* to include a broader spectrum of *AvrPm3*$^{a2/f2}$ alleles would be an effective strategy to protect against more pathogen races in the US, for example. However, this strategy is not likely to be effective in Europe, where suppression of *Pm3a/f* recognition by *SvrPm3*$^{a1/f1}$ might be effective for all *Avr* alleles, no matter the alterations to specificity. An alternative strategy for Europe might be to alter the specificity of *Pm2*, which recognizes *AvrPm2*, an effector that is structurally related to *SvrPm3*$^{a1/f1}$. Both *AvrPm2* and *SvrPm3*$^{a1/f1}$ share structural similarity to ribonucleases (Praz et al, 2017; Bourras et al, 2015 [Chapter 3]), and therefore belong to the superfamily of RNase-Like Proteins associated with Haustoria (RALPH) effectors (Pedersen et al, 2012; Spanu, 2017). If it would be possible to engineer *Pm2* to recognize *SvrPm3*$^{a1/f1}$, then it could be pyramided with *Pm3a/f* and resistance would be restored. However, this strategy would conversely not be effective against US mildew races. It is therefore important to take into account the natural diversity and basis of recognition specificity within individual target populations. Neglecting complexities might compromise the usefulness of targeted breeding research strategies.

6.7 Inducing arms-race dynamics in a naturally balanced interaction

Early interactions of host and pathogen factors which are not yet polymorphic undergo intense selection pressure, resulting in sequence diversity that, given antagonistic selection pressures, might be maintained by balancing selection (Van der Hoorn et al, 2002). This can happen when multiple alleles of such co-evolving host and pathogen factors exist in both host and pathogen populations which are large and stable (Van der Hoorn et al, 2002; Oliva et al, 2015). Interactions which maintain allelic diversity under balancing selection might be forced into arms-race dynamics when diversity has been reduced and selection pressure heightened. Such is the case we observe for *AvrPm3*$^{a2/f2}$ and its relatively recent expansion along with wheat cultivation in the US. With conflicting mechanisms of co-evolutionary dynamics, it is difficult to associate arms-race or trench warfare (indicative of balancing selection) to the mechanism of interaction between R and AVR proteins, as has been suggested for other interactions for which molecular evidence is lacking (Rose et al, 2004; Barrett et al, 2009; Jia et al, 2016). Based on our results, we propose that co-evolutionary dynamics can be fluid and differ between pathogen populations based on host and pathogen demographics, supporting the observation that studying natural diversity and co-evolution in natural populations and comparing it with observations in agricultural settings is a highly informative strategy for studies of host-pathogen interactions.

References

Abascal F, Zardoya R, Telford MJ. 2010. TranslatorX: multiple alignment of nucleotide sequences guided by amino acid translations. *Nucleic Acids Research* **38:** W7-W13.

Ahmed AA, Pedersen C, Schultz-Larsen T, Kwaaitaal M, Jørgensen HJ, Thordal-Christensen H. 2015. The barley powdery mildew candidate secreted effector protein CSEP0105 inhibits the chaperone activity of a small heat shock protein. *Plant Physiology* **168:** 321-33.

Allen RL, Bittner-Eddy PD, Grenville-Briggs LJ, Meitz JC, Rehmany AP, Rose LE, Beynon JL. 2004. Host-parasite coevolutionary conflict between Arabidopsis and downy mildew. *Science* **306:** 1957–1960.

Allen RL, Meitz JC, Baumber RE, Hall SA, Lee SC, Rose LE, Beynon JL. 2008. Natural variation reveals key amino acids in a downy mildew effector that alters recognition specificity by an *Arabidopsis* resistance gene. *Molecular Plant Pathology* **9:** 511-523.

Aluko G, Martinez C, Tohme J, Castano C, Bergman C, Oard JH. 2004. QTL mapping of grain quality traits from the interspecific cross *Oryza sativa* x *O. glaberrima*. *Theoretical Applied Genetics* **109:** 630–639.

Anderson PA, Lawrence GJ, Morrish BC, Ayliffe MA, Finnegan EJ, Ellis JG. 1997. Inactivation of the flax rust resistance gene *M* associated with loss of a repeated unit within the leucine-rich repeat coding region. *Plant Cell* **9:** 641–651.

Aranko AS, Oeemig JS, Kajander T, Iwaï H. 2013. Intermolecular domain swapping induces intein-mediated protein alternative splicing. *Nature Chemical Biology* **9:** 616-622.

Assaad FF, Qiu JL, Youngs H, Ehrhardt D, Zimmerli L, Kalde M, Wanner G, Peck SC, Edwards H, Ramonell K, Somerville CR. 2004. The PEN1 syntaxin defines a novel cellular compartment upon fungal attack and is required for the timely assembly of papillae. *Molecular Biology of the Cell* **15:** 5118-5129.

Bachmann AL, Mootz HD. 2015. An unprecedented combination of serine and cysteine nucleophiles in a split intein with an atypical split site. *Journal of Biological Chemistry* **290:** 28792-28804.

Bakkeren G, Valent B. 2014. Do pathogen effectors play peek-a-boo? *Frontiers in Plant Science* **5:** 731.

Balesdent MH, Attard A, Kuhn ML, Rouxel T. 2002. New avirulence genes in the phytopathogenic fungus *Leptosphaeria maculans*. *Phytopathology* **92:** 1122–33.

Bailey TL, Johnson J, Grant CE, Noble WS. 2015. The MEME suite. *Nucleic Acids Research* **43:** W39-49.

Barrett LG, Thrall PH, Burdon JJ, Linde CC. 2008. Life history determines genetic structure and evolutionary potential of host–parasite interactions. *Trends in Ecology & Evolution* **23:** 678-685.

Barrett LG, Thrall PH, Dodds PN, Van der Merwe M, Linde CC, Lawrence GJ, Burdon JJ. 2009. Diversity and evolution of effector loci in natural populations of the plant pathogen *Melampsora lini*. *Molecular Biology and Evolution* **26:** 2499-2513.

Barrett L, Encinas-Viso F, Thrall PH, Burdon J. 2015. Specialization for resistance in wild host-pathogen interaction networks. *Frontiers in Plant Science* **6:** 761.

Bartsch M, Gobbato E, Bednarek P, Debey S, Schultze JL, Bautor J, Parker JE. 2006. Salicylic acid-independent ENHANCED DISEASE SUSCEPTIBILITY1 signaling in *Arabidopsis*

immunity and cell death is regulated by the monooxygenase FMO1 and the nudix hydrolase NUDT7. *The Plant Cell* **18:** 1038-1051.

Batagelj V, Mrvar A. 2003. *Graph Drawing Software: Pajek- Analysis and Visualization of Large Networks*. Berlin, Germany: Springer 77-103.

Baxter L, Tripathy S, Ishaque N, Boot N, Cabral A, Kemen E, Thines M, Ah-Fong A, Anderson R, Badejoko W, Bittner-Eddy P, et al. 2010. Signatures of adaptation to obligate biotrophy in the *Hyaloperonospora arabidopsidis* genome. *Science* **30:** 1549-1551.

Bednarek P, Piślewska-Bednarek M, Svatoš A, Schneider B, Doubský J, Mansurova M, Humphry M, Consonni C, Panstruga R, Sanchez-Vallet A, Molina A, Schulze-Lefert P. 2009. A glucosinolate metabolism pathway in living plant cells mediates broad-spectrum antifungal defense. *Science* **323:** 101-106.

Ben-David R, Parks R, Dinoor A, Kosman E, Wicker T, Keller B, Cowger C. 2016. Differentiation Among *Blumeria graminis* f. sp. *tritici* Isolates Originating from Wild Versus Domesticated *Triticum* Species in Israel. *Phytopathology* **106:** 861-870.

Beyer A, Bandyopadhyay S, Ideker T. 2007. Integrating physical and genetic maps: from genomes to interaction networks. *Nature Reviews Genetics* **8:** 699-710.

Bhat RA, Miklis M, Schmelzer E, Schulze-Lefert P, Panstruga R. 2005. Recruitment and interaction dynamics of plant penetration resistance components in a plasma membrane microdomain. *Proceedings of the National Academy of Sciences* **102:** 3135-3140.

Bhullar NK, Street K, Mackay M, Yahiaoui N, Keller B. 2009. Unlocking wheat genetic resources for the molecular identification of previously undescribed functional alleles at the *Pm3* resistance locus. *Proceedings of the National Academy of Science* **106:** 9519–9524.

Bhullar NK, Zhang Z, Wicker T, Keller B. 2010. Wheat gene bank accessions as a source of new alleles of the powdery mildew resistance gene *Pm3*: a large scale allele mining project. *BMC Plant Biology* **10:** 88.

Blanquart F, Kaltz O, Nuismer SL, Gandon S. 2013. A practical guide to measuring local adaptation. *Ecology Letters* **16:** 1195-1205.

Blondeau K, Blaise F, Graille M, Kale SD, Linglin J, Ollivier B, Labarde A, Lazar N, Daverdin G, Balesdent MH, Choi DH, Tyler BM, Rouxel T, van Tilbeurgh H, Fudal Isabelle. 2015. Crystal structure of the effector AvrLm4–7 of Leptosphaeria maculans reveals insights into its translocation into plant cells and recognition by resistance proteins. *The Plant Journal* **83:** 610-624.

Boehnert HU, Fudal I, Dioh W, Tharreau D, Notteghem JL, Lebrun MH. 2004. A putative polyketide synthase/peptide synthetase from *Magnaporthe grisea* signals pathogen attack to resistant rice. *The Plant Cell* **16:** 2499-2513.

Bokor AA, Kohn LM, Poulter RT, van Kan JA. 2012. PRP8 inteins in species of the genus *Botrytis* and other ascomycetes. *Fungal Genetics and Biology* **49:** 250-261.

Boller T, He SY. 2009. Innate immunity in plants: an arms race between pattern recognition receptors in plants and effectors in microbial pathogens. *Science* **324:** 742-744.

Borhan MH, Gunn N, Cooper A, Gulden S, Tör M, Rimmer SR, Holub EB. 2008. WRR4 encodes a TIR-NB-LRR protein that confers broad-spectrum white rust resistance in *Arabidopsis thaliana* to four physiological races of *Albugo candida*. *Molecular Plant-Microbe Interactions* **21:** 757-768.

Borhan MH, Holub EB, Kindrachuk C, Omidi M, Bozorgmanesh-Frad G, Rimmer SR. 2010. WRR4, a broad-spectrum TIR-NB-LRR gene from *Arabidopsis thaliana* that confers white rust resistance in transgenic oilseed brassica crops. *Molecular Plant Pathology* **11:** 283-291.

Bos JI, Armstrong MR, Gilroy EM, Boevink PC, Hein I, Taylor RM, Zhendong T, Engelhardt S, Vetukuri RR, Harrower B, Dixelius C, Bryan G, Sadanandom A, Whisson SC, Kamoun S, Birch PRJ. 2010. *Phytophthora infestans* effector AVR3a is essential for virulence and manipulates plant immunity by stabilizing host E3 ligase CMPG1. *Proceedings of the National Academy of Sciences* **107:** 9909-9914.

Bourras S, McNally KE, Ben-David R, Parlange F, Roffler S, Praz CR, Oberhaensli S, Menardo F, Stirnweis D, Frenkel Z, Schaefer LK, Flükiger S, Treier G, Herren G, Korol AB, Wicker T, Keller B. 2015. Multiple avirulence loci and allele-specific effector recognition control the Pm3 race-specific resistance of wheat to powdery mildew. *The Plant Cell* **27:** 2991-3012.

Bourras S, McNally KE, Müller MC, Wicker T, Keller B. 2016. Avirulence genes in cereal powdery mildews: The gene-for-gene hypothesis 2.0. *Frontiers in Plant Science* **7:** 241.

Bousset L, Chèvre AM. 2013. Stable epidemic control in crops based on evolutionary principles: adjusting the metapopulation concept to agro-ecosystems. *Agriculture, Ecosystems & Environment* **165:** 118-129.

Braun U. 2011. The current systematics and taxonomy of the powdery mildews (*Erysiphales*): an overview. *Mycoscience* **52:** 210–212.

Brown JKM, Wolfe MS. 1990. Structure and evolution of a population of *Erysiphe graminis* f. sp. *hordei*. *Plant Pathology* **39:** 376-390.

Brown JK, Simpson CG. 1994. Genetic analysis of DNA fingerprints and virulences in *Erysiphe graminis* f.sp. *hordei*. *Current Genetics* **26:** 172-178.

Brown JKM, Jessop AC. 1995. Genetics of avirulences in *Erysiphe graminis* f.sp. *hordei*. *Plant Pathology* **44:** 1039-1049.

Brown JKM, Le Boulaire S, Evans N. 1996. Genetics of responses to morpholine-type fungicides and of avirulences in *Erysiphe graminis* f. sp. *hordei*. *European Journal of Plant Pathology* **102:** 479-490.

Brunner S, Hurni S, Streckeisen P, Mayr G, Albrecht M, Yahiaoui N, Keller B. 2010. Intragenic allele pyramiding combines different specificities of wheat *Pm3* resistance alleles. *The Plant Journal* **64:** 433-445.

Bull JJ. 1994. Perspective: virulence. *Evolution* **1:** 1423-1437.

Burdon JJ, Thrall PH. 2009. Coevolution of plants and their pathogens in natural habitats. *Science* **324:** 755-756.

Bürglin TR. 2008. The Hedgehog protein family. *Genome Biology* **9:** 241.

Bustin SA, Benes V, Garson JA, Hellemans J, Huggett J, Kubista M, Mueller R, Nolan T, Pfaffl MW, Shipley GL, Vandesompele J, Wittwer CT. 2009. The MIQE guidelines: minimum information for publication of quantitative real-time PCR experiments. *Clinical Chemistry* **55:** 611-622.

Caffier V, de Vallavieille-Pope C, Brown JKM. 1996. Segregation of avirulences and genetic basis of infection types in *Erysiphe graminis* f.sp. *hordei*. *Phytopathology* **86:** 1112-1121.

Catanzariti AM, Dodds PN, Ve T, Kobe B, Ellis JG, Staskawicz BJ. 2010. The *AvrM* effector from flax rust has a structured C-terminal domain and interacts directly with the *M* resistance protein. *Molecular Plant Microbe Interactions* **23:** 49-57.

Cesari S, Thilliez G, Ribot C, Chalvon V, Michel C, Jauneau A, Rivas S, Alaux L, Kanzaki H, Okuyama Y, Morel JB, Fournier E, Tharreau D, Terauchi R, Kroj T. 2013. The rice resistance protein pair *RGA4/RGA5* recognizes the *Magnaporthe oryzae* effectors *AVR-Pia* and *AVR1-CO39* by direct binding. *The Plant Cell* **25:** 1463-1481.

Chen X, Yang JY. 2010. Constructing consensus genetic maps in comparative analysis. *Journal of Computational Biology* **17:** 1561-1573.

Chen XL, Shi T, Yang J, Shi W, Gao X, Chen D, Xu X, Xu JR, Talbot NJ, Peng YL. 2014. N-glycosylation of effector proteins by an α-1, 3-mannosyltransferase is required for the rice blast fungus to evade host innate immunity. *The Plant Cell* **26:** 1360-1376.

Chou S, Krasileva KV, Holton JM, Steinbrenner AD, Alber T, Staskawicz BJ. 2011. *Hyaloperonospora arabidopsidis* ATR1 effector is a repeat protein with distributed recognition surfaces. *Proceedings of the National Academy of Sciences* **108:** 13323-13328.

Collins NC, Thordal-Christensen H, Lipka V, Bau S, Kombrink E, Qiu JL, Hückelhoven R, Stein M, Freialdenhoven A, Somerville SC, Schulze-Lefert P. 2003. SNARE-protein-mediated disease resistance at the plant cell wall. *Nature* **425:** 973-977.

Consonni C, Humphry ME, Hartmann HA, Livaja M, Durner J, Westphal L, Vogel J, Lipka V, Kemmerling B, Schulze-Lefert P, Somerville SC, Panstruga R. 2006. Conserved requirement for a plant host cell protein in powdery mildew pathogenesis. *Nature Genetics* **38:** 716-720.

Cowger C, Mehra L, Arellano C, Meyers E, Murphy JP. *in review.* Virulence differences in *Blumeria graminis* f. sp. *tritici* from the central and eastern United States. *Phytopathology.*

Cowger C, Parks R, Kosman E. 2016. Structure and migration in US *Blumeria graminis* f. sp. *tritici* populations. *Phytopathology* **106:** 295-304.

Croll D, McDonald BA. 2012. The accessory genome as a cradle for adaptive evolution in pathogens. *PLoS Pathogens* **8:** e1002608.

Dangl JL, McDowell JM. 2006. Two modes of pathogen recognition by plants. *Proceedings of the National Academy of Sciences* **103:** 8575-8576.

Daverdin G, Rouxel T, Gout L, Aubertot JN, Fudal I, Meyer M, Parlange F, Carpezat J, Balesdent MH. 2012. Genome structure and reproductive behaviour influence the evolutionary potential of a fungal phytopathogen. *PLoS Pathogens* **8:** e1003020.

de Guillen K, Ortiz-Vallejo D, Gracy J, Fournier E, Kroj T, Padilla A. 2015. Structure analysis uncovers a highly diverse but structurally conserved effector family in phytopathogenic fungi. *PLoS Pathogens* **11:** e1005228.

de Jonge R, van Esse HP, Kombrink A, Shinya T, Desaki Y, Bours R, van der Krol S, Shibuya N, Joosten MH, Thomma BP. 2010. Conserved fungal LysM effector Ecp6 prevents chitin-triggered immunity in plants. *Science* **329:** 953–955.

Dido AA, San FY, Niks RE. 2016. Evaluation of plant stage dependency of QTLs to homologous and heterologous rust pathogen isolates of barley. *African Journal of Plant Science*, **10:** 130-135.

Dodds PN, Lawrence GJ, Catanzariti A, Ayliffe MA, Ellis JG. 2004. The *Melampsora lini AvrL567* avirulence genes are expressed in haustoria and their products are recognized inside plant cells. *The Plant Cell* **16:** 755–768.

Dodds PN, Lawrence GJ, Catanzariti AM, Teh T, Wang CI, Ayliffe MA, Kobe B, Ellis JG. 2006. Direct protein interaction underlies gene-for-gene specificity and coevolution of the flax resistance genes and flax rust avirulence genes. *Proceedings of the National Academy of Sciences* **103:** 8888-8893.

Dodds PN, Rafiqi M, Gan PH, Hardham AR, Jones DA, Ellis JG. 2009. Effectors of biotrophic fungi and oomycetes: pathogenicity factors and triggers of host resistance. *New Phytologist* **183:** 993-1000.

Drozdetskiy A, Cole C, Procter J, Barton GJ. 2015. JPred4: a protein secondary structure prediction server. *Nucleic Acids Research* **16:** gkv332.

Dugdale B, Mortimer CL, Kato M, James TA, Harding RM, Dale JL. 2014. Design and construction of an in-plant activation cassette for transgene expression and recombinant protein production in plants. *Nature Protocols* **9:** 1010–1027.

Edgar RC. 2004. MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Research* **32:** 1792-1797.

Eisenach C, Chen ZH, Grefen C, Blatt MR. 2012. The trafficking protein SYP121 of *Arabidopsis* connects programmed stomatal closure and K(+) channel activity with vegetative growth. *Plant Journal* **69:** 241–251.

Ellis JG, Lawrence GJ, Luck JE, Dodds PN. 1999. Identification of regions in alleles of the flax rust resistance gene *L* that determine differences in gene-for-gene specificity. *The Plant Cell* **11:** 495-506.

Ellis JG, Dodds PN, Lawrence GJ. 2007. Flax rust resistance gene specificity is based on direct resistance-avirulence protein interactions. *Annual Review of Phytopathology* **45:** 289-306.

Eryilmaz E, Shah NH, Muir TW, Cowburn D. 2014. Structural and dynamical features of inteins and implications on protein splicing. *Journal of Biological Chemistry* **289:** 14506-14511.

Eshed NA, Wahl I. 1970. Host ranges and interrelations of *Erysiphe graminis hordei, E. graminis tritici,* and *E. graminis avenae*. *Phytopathology* **60:** 628-634.

Feys BJ, Moisan LJ, Newman MA, Parker JE. 2001. Direct interaction between the *Arabidopsis* disease resistance signaling proteins, EDS1 and PAD4. *The EMBO Journal* **20:** 5400-5411.

Feys BJ, Wiermer M, Bhat RA, Moisan LJ, Medina-Escobar N, Neu C, Cabral A, Parker JE. 2005. *Arabidopsis* SENESCENCE-ASSOCIATED GENE101 stabilizes and signals within an ENHANCED DISEASE SUSCEPTIBILITY1 complex in plant innate immunity. *The Plant Cell* **17:** 2601-2613.

Flor HH. 1956. The complementary genic systems in flax and flax rust. *Advanced Genetics* **8:**29-54.

Flor HH. 1971. Current status of the gene-for-gene concept. *Annual Review of Phytopathology* **9:** 275-296.

Frenkel Z, Paux E, Mester D, Feuillet C, Korol A. 2010. LTC: a novel algorithm to improve the efficiency of contig assembly for physical mapping in complex genomes. *BMC Bioinformatics* **11:** 584.

Fudal I, Ross S, Gout L, Blaise F, Kuhn ML, Eckert MR, Cattolico L, Bernard-Samain S, Balesdent MH, Rouxel T. 2007. Heterochromatin-like regions as ecological niches for avirulence genes in the *Leptosphaeria maculans* genome: map-based cloning of *AvrLm6*. *Molecular Plant-Microbe Interactions* **20:** 459–70.

Galon Y, Nave R, Boyce JM, Nachmias D, Knight MR, Fromm H. 2008. Calmodulin-binding transcription activator (CAMTA) 3 mediates biotic defense responses in *Arabidopsis*. *FEBS Letters* **582:** 943-948.

Gawehns F, Houterman PM, Ichou FA, Michielse CB, Hijdra M, Cornelissen BJC, Rep M, Takken FLW. 2014. The *Fusarium oxysporum* effector *Six6* contributes to virulence and suppresses *I-2* mediated cell death. *Molecular Plant Microbe Interactions* **27:** 336–48.

Ghanbarnia K, Fudal I, Larkan NJ, Links MG, Balesdent MH, Profotova B, Fernando WG, Rouxel T, Borhan MH. 2015. Rapid identification of the *Leptosphaeria maculans* avirulence gene *AvrLm2* using an intraspecific comparative genomics approach. *Molecular Plant Pathology* **16:** 699-709.

Gijzen M, Ishmael C, Shrestha SD. 2014. Epigenetic control of effectors in plant pathogens. *Frontiers in Plant Sciences* **5:** 1–4.

Gill US, Lee S, Mysore KS. 2015. Host versus nonhost resistance: distinct wars with similar arsenals. *Phytopathology* **105:** 580-587.

Giraldo MC, Valent B. 2013. Filamentous plant pathogen effectors in action. *Nature Reviews Microbiology* **11:** 800-814.

Godfrey D, Böhlenius H, Pedersen C, Zhang Z, Emmersen J, Thordal-Christensen H. 2010. Powdery mildew fungal effector candidates share N-terminal Y/F/WxC-motif. *BMC Genomics* **11:** 317.

Grandaubert J, Lowe RG, Soyer JL, Schoch CL, Van de Wouw AP, Fudal I, Robbertse B, Lapalu N, Links MG, Ollivier B, Linglin J. 2014. Transposable element-assisted evolution and adaptation to host plant within the *Leptosphaeria maculans-Leptosphaeria biglobosa* species complex of fungal pathogens. *BMC Genomics* **15:** 891.

Hall SA, Allen RL, Baumber RE, Baxter LA, Fisher K, Bittner-Eddy PD, Rose LE, Holub EB, Beynon JL. 2009. Maintenance of genetic variation in plants and pathogens involves complex networks of gene-for-gene interactions. *Molecular Plant Pathology* **10:** 449–457.

He C, Holme J, Anthony J. 2014. SNP genotyping: the KASP assay. *Methods in Molecular Biology* **45:** 75-86.

Hermansen JE, Torp U, Prahm LP. 1978. Studies of transport of live spores of cereal mildew and rust fungi across the North Sea. *Grana* **17:** 41-46.

Himmelbach A, Zierold U, Hensel G, Riechen J, Douchkov D, Schweizer P, Kumlehn J. 2007. A set of modular binary vectors for transformation of cereals. *Plant Physiology* **145**:1192-1200.

Houterman PM, Cornelissen BJC, Rep M. 2008. Suppression of plant resistance gene-based immunity by a fungal effector. *PLoS Pathogens* **4:** e1000061.

Howard RJ, Valent B. 1996. Breaking and entering: host penetration by the fungal rice blast pathogen *Magnaporthe grisea*. *Annual Reviews in Microbiology* **50:** 491-512.

Howles P, Lawrence G, Finnegan J, McFadden H, Ayliffe M, Dodds P, Ellis J. 2005. Autoactive alleles of the flax *L6* rust resistance gene induce non-race-specific rust resistance associated with the hypersensitive response. *Molecular Plant Microbe Interactions* **18:** 570-582.

Huang YJ, Li ZQ, Evans N, Rouxel T, Fitt BD, Balesdent MH. 2006. Fitness cost associated with loss of the *AvrLm4* avirulence function in *Leptosphaeria maculans* (phoma stem canker of oilseed rape). *European Journal of Plant Pathology* **114:** 77-89.

Huang YJ, Balesdent MH, Li ZQ, Evans N, Rouxel T, Fitt BD. 2010. Fitness cost of virulence differs between the *AvrLm1* and *AvrLm4* loci in *Leptosphaeria maculans* (phoma stem canker of oilseed rape). *European Journal of Plant Pathology* **126:** 279.

Huang J, Si W, Deng Q, Li P, Yang S. 2014. Rapid evolution of avirulence genes in rice blast fungus *Magnaporthe oryzae*. *BMC Genetics* **15:** 45.

Inami K, Yoshioka-Akiyama C, Morita Y, Yamasaki M, Teraoka T, Arie T. 2012. A genetic mechanism for emergence of races in *Fusarium oxysporum* f. sp. *lycopersici*: inactivation of avirulence gene *AVR1* by transposon insertion. *PLos One* **7:** e44101.

Jacobson ES. 2000. Pathogenic roles for fungal melanins. *Clinical Microbiology Reviews* **13:** 708-717.

Jafary H, Albertazzi G, Marcel TC, Niks RE. 2008. High diversity of genes for nonhost resistance of barley to heterologous rust fungi. *Genetics* **178:** 2327-2339.

Jia Y, McAdams SA, Bryan GT, Hershey HP, Valent B. 2000. Direct interaction of resistance gene and avirulence gene products confers rice blast resistance. *The EMBO Journal* **19:** 4004-4014.

Jia Y, Zhou E, Lee S, Bianco T. 2016. Coevolutionary dynamics of rice blast resistance gene *Pi-ta* and *Magnaporthe oryzae* avirulence gene *AVR-Pita 1*. *Phytopathology* **106:** 676-683.

Johansson ON, Fantozzi E, Fahlberg P, Nilsson AK, Buhot N, Tör M, Andersson MX. 2014. Role of the penetration-resistance genes PEN1, PEN2 and PEN3 in the hypersensitive response and race-specific resistance in *Arabidopsis thaliana*. *The Plant Journal* **79:** 466-476.

Johnson R. 1984. A critical analysis of durable resistance. *Annual Review of Phytopathology* **22:** 309-330.

Jones DA. 1988a. Genes for resistance to flax rust in the flax cultivars Towner and Victory And the genetics of pathogenicity in flax rust to the *L8* gene for resistance. *Phytopathology* **78:** 338–41.

Jones DA. 1988b. Genetic properties on inhibitor genes in flax rust that alter avirulence to virulence on flax. *Phytopathology* **78:** 342–44.

Jones JDG, Dangl JL. 2006. The plant immune system. *Nature* **444:** 323–329.

Joosten MH, de Wit PJ. 1999. The Tomato–C ladosporium F ulvum Interaction: A Versatile Experimental System to Study Plant-Pathogen Interactions. *Annual Review of Phytopathology* **37:** 335-367.

Jordan T, Draeger L, Kuhn BM, Pedersen C, Anderson MX, Maekawa T, Thordal-Christensen H, Schulze-Lefert P, Keller B. Powdery mildew formae speciales and isolates display different levels of virulence on the non-host plant Arabidopsis. *In preparation*.

Källberg M, Margaryan G, Wang S, Ma J, Xu J. 2014. RaptorX server: a resource for template-based protein structure modeling. In: Kihara D, eds. *Protein Structure Prediction*. New York, USA: Springer New York, 17-27.

Kanzaki H, Yoshida K, Saitoh H, Fujisaki K, Hirabuchi A, Alaux L, Fournier E, Tharreau D, Terauchi R. 2012. Arms race co-evolution of *Magnaporthe oryzae AVR-Pik* and rice *Pik* genes driven by their physical interactions. *The Plant Journal* **72:** 894-907.

Kelley LA, Mezulis S, Yates CM, Wass MN, Sternberg MJ. 2015. The Phyre2 web portal for protein modeling, prediction and analysis. *Nature Protocols* **10:** 845-858.

Kobae Y, Sekino T, Yoshioka H, Nakagawa T, Martinoia E, Maeshima M. 2006. Loss of AtPDR8, a plasma membrane ABC transporter of *Arabidopsis thaliana*, causes hypersensitive cell death upon pathogen infection. *Plant and Cell Physiology* **47:** 309-318.

Koeck M, Hardham AR, Dodds PN. 2011. The role of effectors of biotrophic and hemibiotrophic fungi in infection. *Cellular Microbiology* **13:** 1849-1857.

Krasileva KV, Dahlbeck D, Staskawicz BJ. 2010. Activation of an Arabidopsis resistance protein is specified by the *in planta* association of its leucine-rich repeat domain with the cognate oomycete effector. *The Plant Cell* **22:** 2444-2458.

Krattinger SG, Keller B. 2016. Molecular genetics and evolution of disease resistance in cereals. *New Phytologist* **212:** 320-332.

Krupinsky JM, Bailey KL, McMullen MP, Gossen BD, Turkington TK. 2002. Managing plant disease risk in diversified cropping systems. *Agronomy Journal* **94:** 198-209.

Kumar S, Stecher G, Tamura K. 2016. MEGA7: Molecular Evolutionary Genetics Analysis version 7.0 for bigger datasets. *Molecular Biology and Evolution* **33:** 1870-1874.

Kwaaitaal M, Keinath NF, Pajonk S, Biskup C, Panstruga R. 2010. Combined bimolecular fluorescence complementation and Förster resonance energy transfer reveals ternary SNARE complex formation in living plant cells. *Plant Physiology* **152:** 1135-1147.

Kwon C, Neu C, Pajonk S, Yun HS, Lipka U, Humphry M, Bau S, Straus M, Kwaaitaal M, Rampelt H, El Kasmi F. 2008. Co-option of a default secretory pathway for plant immune responses. *Nature* **451:** 835-840.

Laine AL. 2005. Spatial scale of local adaptation in a plant-pathogen metapopulation. *Journal of Evolutionary Biology* **18:** 930-938.

Laine AL, Burdon JJ, Dodds PN, Thrall PH. 2011. Spatial variation in disease resistance: from molecules to metapopulations. *Journal of Ecology* **99:** 96-112.

Laine AL, Burdon JJ, Nemri A, Thrall PH. 2014. Host ecotype generates evolutionary and epidemiological divergence across a pathogen metapopulation. *Proceedings of the Royal Society of London B: Biological Sciences* **281:** 20140522.

Lander ES, Green P, Abrahamson J, Barlow A, Daly MJ, Lincoln SE, Newberg LA, Newburg L. 1987. MAPMAKER: an interactive computer package for constructing primary genetic linkage maps of experimental and natural populations. *Genomics* **1:** 174–181.

Larkin MA, Blackshields G, Brown NP, Chenna R, McGettigan PA, McWilliam H, Valentin F, Wallace IM, Wilm A, Lopez R, Thompson JD, Gibson TJ, Higgins DG. 2007. Clustal W and Clustal X version 2.0. *Bioinformatics* **23:** 2947-2948.

Lawrence GJ, Mayo GME, Shepherd KW. 1981. Interactions Between Genes Controlling Pathogenicity in the Flax Rust Fungus. *Phytopathology* **71:** 12-19.

Leigh JW, Bryant D. 2015. Popart: full-feature software for haplotype network construction. *Methods in Ecology and Evolution* **6:** 1110-1116.

Lendenmann MH, Croll D, Stewart EL, McDonald BA. 2014. Quantitative trait locus mapping of melanization in the plant pathogenic fungus *Zymoseptoria tritici*. *G3: Genes, Genomes, Genetics* **4:** 2519-2533.

Leonelli L, Pelton J, Schoeffler A, Dahlbeck D, Berger J, Wemmer DE, Staskawicz B. 2011. Structural elucidation and functional characterization of the *Hyaloperonospora arabidopsidis* effector protein *ATR13. PLoS Pathogens* **7:** e1002428.

Lévesque CA, Brouwer H, Cano L, Hamilton JP, Holt C,Huitema E, Raffaele S, Robideau GP, Thines M, Win J, Zerillo MM, Beakes GW, Boore JL, Busam D, Dumas B, Ferriera S, Fuerstenberg SI, Gachon CMM, Gaulin E, Govers F, Grenville-Briggs L, Horner N, Hostetler J, Jiang RHY, Johnson J, Krajaejun T, Lin H, Meijer HJG, Moore B, Morris P, Phuntmart V, Puiu D, Shetty J, Stajich JE, Tripathy S, Wawra S, vanWest P, Whitty BR, Coutinho PM, Henrissat B, Martin F, Thomas PD, Tyler BM, De Vries RP, Kamoun S, Yandell M, Tisserat N, Buell CR. 2010. Genome sequence of the necrotrophic plant pathogen Pythium ultimum reveals original pathogenicity mechanisms and effector repertoire. *Genome Biology* **11:** R73.

Li X, Kapos P, Zhang Y. 2015. NLRs in plants. *Current Opinion in Immunology* **32:** 114-121.

Lipka V, Dittgen J, Bednarek P, Bhat R, Wiermer M, Stein M, Landtag J, Brandt W, Rosahl S, Scheel D, Llorente F, et al. 2005. Pre-and postinvasion defenses both contribute to nonhost resistance in *Arabidopsis*. *Science* **310:** 1180-1183.

Lipka AE, Tian F, Wang Q, Peiffer J, Li M, Bradbury PJ, Gore MA, Buckler ES, Zhang Z. 2012. GAPIT: genome association and prediction integrated tool. *Bioinformatics* **28:** 2397-2399.

Lo Presti L, Lanver D, Schweizer G, Tanaka S, Liang L, Tollot M, Zuccaro A, Reissmann S, Kahmann R. 2015. Fungal effectors and plant susceptibility. *Annual Review of Plant Biology* **66:** 513-545.

Lu X, Kracher B, Saur IM, Bauer S, Ellwood SR, Wise R, Yaeno T, Maekawa T, Schulze-Lefert P. 2016. Allelic barley MLA immune receptors recognize sequence-unrelated avirulence effectors of the powdery mildew pathogen. *Proceedings of the National Academy of Sciences* **4:** 201612947.

Ma L, Lukasik E, Gawehns F, Takken FL. 2012. The use of agroinfiltration for transient expression of plant resistance and fungal effector proteins in Nicotiana benthamiana leaves. In: Bolton MD, Thomma BPHJ, eds. *Plant Fungal Pathogens: Methods and Protocols*. New York, USA: Humana Press, 61–74.

Ma L, Houterman PM, Gawehns F, Cao L, Sillo F, Richter H, Clavijo-Ortiz MJ, Schmidt SM, Boeren S, Vervoort J, Cornelissen BJC, Rep M, Takken FLW. 2015. The *AVR2 – SIX5* gene pair is required to activate *I-2* -mediated immunity in tomato. *New Phytologist* **doi:** 10.1111/nph.13455.

Maekawa T, Kufer TA, Schulze-Lefert P. 2011. NLR functions in plant and animal immune systems: so far and yet so close. *Nature Immunology* **12:** 817-826.

Malinovsky FG, Fangel JU, Willats WG. 2014. The role of the cell wall in plant immunity. *Frontiers in Plant Science* **5:** 38-49.

Maqbool A, Saitoh H, Franceschetti M, Stevenson CEM, Uemura A, Kanzaki H, Kamoun S, Terauchi R, Banfield MJ. 2015. Structural basis of pathogen recognition by an integrated HMA domain in a plant NLR immune receptor. *Elife* **4:** e08709.

Marone D, Russo MA, Laidò G, De Leonardis AM, Mastrangelo AM. 2013. Plant nucleotide binding site–leucine-rich repeat (NBS-LRR) genes: active guardians in host defense responses. *International Journal of Molecular Sciences* **14:** 7302-7326.

Matsumura H, Takahashi H, Inoue T, Yamamoto T, Hashimoto H, Nishioka M, Fujiwara S, Takagi M, Imanaka T, Kai Y. 2006. Crystal structure of intein homing endonuclease II encoded in DNA polymerase gene from hyperthermophilic archaeon *Thermococcus kodakaraensis* strain KOD1. *Proteins: Structure, Function, and Bioinformatics* **63:** 711-715.

McDonald BA, Linde C. 2002. Pathogen population genetics, evolutionary potential, and durable resistance. *Annual Review of Phytopathology* **40:** 349-379.

Menardo F, Praz CR, Wyder S, Ben-David R, Bourras S, Matsumae H, McNally KE, Parlange F, Riba A, Roffler S, Schaefer LK, et al. 2016. Hybridization of powdery mildew strains gives rise to pathogens on novel agricultural crop species. *Nature Genetics* **48:** 201-205.

Menardo F, Wicker T, Keller B. 2017. Reconstructing the Evolutionary History of Powdery Mildew Lineages (*Blumeria graminis*) at Different Evolutionary Time Scales with NGS Data. *Genome Biology and Evolution* **9:** 446.

Meyer D, Pajonk S, Micali C, O'Connell R, Schulze-Lefert P. 2009. Extracellular transport and integration of plant secretory proteins into pathogen-induced cell wall compartments. *The Plant Journal* **57:** 986-999.

Mills KV, Johnson MA, Perler FB. 2014. Protein splicing: how inteins escape from precursor proteins. *Journal of Biological Chemistry* **289:** 14498-14505.

Moffett P, Farnham G, Peart J, Baulcombe DC. 2002. Interaction between domains of a plant NBS-LRR protein in disease resistance-related cell death. *EMBO Journal* **21:** 4511-4519.

Nei M, Kumar S. 2000. *Molecular Evolution and Phylogenetics*. New York, USA: Oxford University Press.

Nemri A, Saunders DG, Anderson C, Upadhyaya NM, Win J, Lawrence G, Jones D, Kamoun S, Ellis J, Dodds P. 2014. The genome sequence and effector complement of the flax rust pathogen *Melampsora lini*. *Frontiers in Plant Science* **5:** 98.

Niemi L, Wennström A, Hjältén J, Waldmann P, Ericson L. 2006. Spatial variation in resistance and virulence in the host–pathogen system *Salix triandra–Melampsora amygdalinae*. *Journal of Ecology* **94:** 915-921.

Nielsen R, Yang Z. 1998. Likelihood models for detecting positively selected amino acid sites and applications to the *HIV-1* envelope gene. *Genetics* **148:** 929-936.

Nielsen ME, Feechan A, Böhlenius H, Ueda T, Thordal-Christensen H. 2012. *Arabidopsis* ARF-GTP exchange factor, GNOM, mediates transport required for innate immunity and focal accumulation of syntaxin PEN1. *Proceedings of the National Academy of Sciences* **109:** 11443-11448.

Niewoehner AS, Leath S. 1998. Virulence of *Blumeria graminis* f. sp. *tritici* on winter wheat in the eastern United States. *Plant Disease* **82:** 64-68.

Novikova O, Topilina N, Belfort M. 2014. Enigmatic distribution, evolution, and function of inteins. *Journal of Biological Chemistry* **289:** 14490-14497.

Oberhaensli S, Parlange F, Buchmann JP, Jenny FH, Abbott JC, Burgis TA, Spanu PD, Keller B, Wicker T. 2011. Comparative sequence analysis of wheat and barley powdery mildew fungi reveals gene colinearity, dates divergence and indicates host-pathogen co-evolution. *Fungal Genetics and Biology* **48:** 327-334.

Oeemig JS, Aranko AS, Djupsjöbacka J, Heinämäki K, Iwaï H. 2009. Solution structure of DnaE intein from *Nostoc punctiforme*: Structural basis for the design of a new split intein suitable for site-specific chemical modification. *FEBS Letters* **583:** 1451-1456.

Oliva RF, Cano LM, Raffaele S, Win J, Bozkurt TO, Belhaj K, Oh SK, Thines M, Kamoun S. 2015. A recent expansion of the RXLR effector gene *Avrblb2* is maintained in global populations of *Phytophthora infestans* indicating different contributions to virulence. *Molecular Plant-Microbe Interactions* **28:** 901-912.

Ortiz D, De Guillen K, Cesari S, Chalvon V, Gracy J, Padilla A, Kroj T. 2017. Recognition of the Magnaporthe oryzae effector AVR-Pia by the decoy domain of the rice NLR immune receptor RGA5. *The Plant Cell* **29:** 156-68.

Özkan H, Brandolini A, Schäfer-Pregl R, Salamini F. 2002. AFLP analysis of a collection of tetraploid wheats indicates the origin of emmer and hard wheat domestication in southeast Turkey. *Molecular Biology and Evolution* **19:** 1797-1801.

Pajonk S, Kwon C, Clemens N, Panstruga R, Schulze-Lefert P. 2008. Activity determinants and functional specialization of *Arabidopsis* PEN1 syntaxin in innate immunity. *Journal of Biological Chemistry* **283:** 26974-26984.

Parker JE, Holub EB, Frost LN, Falk A, Gunn ND, Daniels MJ. 1996. Characterization of eds1, a mutation in *Arabidopsis* suppressing resistance to *Peronospora parasitica* specified by several different RPP genes. *The Plant Cell* **8:** 2033-2046.

Parlange F, Daverdin G, Fudal I, Kuhn ML, Balesdent MH, Blaise F, Grezes-Besset B, Rouxel T. 2009. *Leptosphaeria maculans* avirulence gene *AvrLm4-7* confers a dual recognition specificity by the *Rlm4* and *Rlm7* resistance genes of oilseed rape, and circumvents *Rlm4*-mediated recognition through a single amino acid change. *Molecular Microbiology* **71:** 851-863.

Parlange F, Oberhaensli S, Breen J, Platzer M, Taudien S, Šimková H, Wicker T, Doležel J, Keller B. 2011. A major invasion of transposable elements accounts for the large size of the *Blumeria graminis* f. sp. *tritici* genome. *Functional & Integrative Genomics* **11:** 671-677.

Parlange F, Roffler S, Menardo F, Ben-David R, Bourras S, McNally KE, Oberhaensli S, Stirnweis D, Buchmann G, Wicker T, Keller B. 2015. Genetic and molecular characterization of a locus involved in avirulence of *Blumeria graminis* f. sp. *tritici* on wheat *Pm3* resistance alleles. *Fungal Genetics and Biology* **82:** 181-192.

Parniske M, Hammond-Kosack KE, Golstein C, Thomas CM, Jones DA, Harrison K, Wulff BB, Jones JD. 1997. Novel disease resistance specificities result from sequence exchange between tandemly repeated genes at the *Cf-4/9* locus of tomato. *Cell* **91:** 821-832.

Pedersen C, van Themaat EV, McGuffin LJ, Abbott JC, Burgis TA, Barton G, Bindschedler LV, Lu X, Maekawa T, Weßling R, Cramer R. 2012. Structure and evolution of barley powdery mildew effector candidates. *BMC Genomics* **13:** 694.

Peng JH, Sun D, Nevo E. 2011. Domestication evolution, genetics and genomics in wheat. *Molecular Breeding* **28:** 281.

Petersen TN, Brunak S, von Heijne G, Nielsen H. 2011. SignalP 4.0: Discriminating signal peptides from transmembrane regions. *Nature Methods* **8:** 785-786.

Pietrokovski S. 2001. Intein spread and extinction in evolution. *TRENDS in Genetics* **17:** 465-472.

Pliego C, Nowara D, Bonciani G, Gheorghe DM, Xu R, Surana P, Whigham E, Nettleton D, Bogdanove AJ, Wise RP, Schweizer P. 2013. Host-induced gene silencing in barley powdery mildew reveals a class of ribonuclease-like effectors. *Molecular Plant-Microbe Interactions* **26:** 633-642.

Pond SL, Frost SD, Muse SV. 2005. HyPhy: hypothesis testing using phylogenies. *Bioinformatics* **21:** 676-9.

Praz CR, Bourras S, Zeng F, Sánchez-Martín J, Menardo F, Xue M, Yang L, Roffler S, Böni R, Herren G, McNally KE, Ben-David R, Parlange F, Oberhaensli S, Flükiger S, Schäfer LK, Wicker T, Keller B. 2017. *AvrPm2* encodes an RNase-like avirulence effector which is conserved in the two different specialized forms of wheat and rye powdery mildew fungus. *New Phytologist* **213:** 1301-1314.

Qi D, Innes RW. 2013. Recent advances in plant NLR structure, function, localization, and signaling. *Frontiers in Immunology* **4:** 348.

Raffaele S, Farrer RA, Cano LM, Studholme DJ, MacLean D, Thines M, Jiang RH, Zody MC, Kunjeti SG, Donofrio NM, Meyers BC, Nusbaum C, Kamoun S. 2010. Genome evolution following host jumps in the Irish potato famine pathogen lineage. *Science* **330:** 1540-1543.

Raffaele S, Kamoun S. 2012. Genome evolution in filamentous plant pathogens: why bigger can be better. *Nature Review Microbiology* **10:** 417–430.

Rafiqi M, Ellis JG, Ludowici VA, Hardham AR, Dodds PN. 2012. Challenges and progress towards understanding the role of effectors in plant-fungal interactions. *Current Opinion in Plant Biology* **15:** 477–82.

Ravensdale M, Nemri A, Thrall PH, Ellis JG, Dodds PN. 2011. Co-evolutionary interactions between host resistance and pathogen effector genes in flax rust disease. *Molecular Plant Pathology* **12:** 93–102.

Ravensdale M, Bernoux M, Ve T, Kobe B, Thrall PH, Ellis JG, Dodds PN. 2012. Intramolecular interaction influences binding of the Flax L5 and L6 resistance proteins to their AvrL567 ligands. *PLoS Pathogens* **8:** e1003004.

Ridout CJ, Skamnioti P, Porritt O, Sacristan S, Jones JDG, Brown JKM. 2006. Multiple avirulence paralogues in cereal powdery mildew fungi may contribute to parasite fitness and defeat of plant resistance. *The Plant Cell* **18:** 2402–2414.

Rietz S, Stamm A, Malonek S, Wagner S, Becker D, Medina-Escobar N, Vlot AC, Feys BJ, Niefind K, Parker JE. 2011. Different roles of Enhanced Disease Susceptibility1 (EDS1) bound to and dissociated from Phytoalexin Deficient4 (PAD4) in *Arabidopsis* immunity. *New Phytologist* **191:** 107-119.

Rooney HC, van't Klooster JW, van der Hoorn RA, Joosten MH, Jones JD, de Wit PJ. 2005. Cladosporium *Avr2* inhibits tomato *Rcr3* protease required for *Cf-2*-dependent disease resistance. *Science* **308:** 1783-1786.

Rose LE, Bittner-Eddy PD, Langley CH, Holub EB, Michelmore RW, Beynon JL. 2004. The maintenance of extreme amino acid diversity at the disease resistance gene, *RPP13*, in *Arabidopsis thaliana*. *Genetics* **166:** 1517-1527.

Rouxel T, Balesdent MH. 2017. Life, death and rebirth of avirulence effectors in a fungal pathogen of Brassica crops, *Leptosphaeria maculans*. *New Phytologist* **214:** 526-532.

Schlaeppi K, Mauch F. 2010. Indolic secondary metabolites protect *Arabidopsis* from the oomycete pathogen *Phytophthora brassicae*. *Plant Signaling & Behavior* **5:** 1099-1101.

Schneider CA, Rasband WS, Eliceiri KW. 2012. NIH Image to ImageJ: 25 years of image analysis. *Nature Methods* **9:** 671.

Schulze-Lefert P, Panstruga R. 2011. A molecular evolutionary concept connecting nonhost resistance, pathogen host range, and pathogen speciation. *Trends in Plant Science* **16:** 117-125.

Schweizer P, Pokorny J, Abderhalden O, Dudler R. 1999. A transient assay system for the functional assessment of defense-related genes in wheat. *Molecular Plant Microbe Interactions* **12:** 647-654.

Seeholzer S, Tsuchimatsu T, Jordan T, Bieri S, Pajonk S, Yang W, Jahoor A, Shimizu KK, Keller B, Schulze-Lefert P. 2010. Diversity at the *Mla* powdery mildew resistance locus from cultivated barley reveals sites of positive selection. *Molecular plant-microbe interactions* **23:** 497-509.

Sela H, Spiridon LN, Ashkenazi H, Bhullar NK, Brunner S, Petrescu AJ, Fahima T, Keller B, Jordan T. 2014. Three-dimensional modeling and diversity analysis reveals distinct AVR recognition sites and evolutionary pathways in wild and domesticated wheat *Pm3 R* genes. *Molecular Plant Microbe Interactions* **27:** 835-845.

Senthil-Kumar M, Mysore KS. 2013. Nonhost resistance against bacterial pathogens: retrospectives and prospects. *Annual Review of Phytopathology* **51:** 407-427.

Shah NH, Eryilmaz E, Cowburn D, Muir TW. 2013. Naturally split inteins assemble through a "capture and collapse" mechanism. *Journal of the American Chemical Society* **135:** 18673-18681.

Shah NH, Muir TW. 2014. Inteins: nature's gift to protein chemists. *Chemical Science* **5:** 446-461.

Slot JC, Hibbett DS. 2007. Horizontal transfer of a nitrate assimilation gene cluster and ecological transitions in fungi: a phylogenetic study. *PloS One* **2:** e1097.

Sohn KH, Lei R, Nemri A, Jones JD. 2007. The downy mildew effector proteins ATR1 and ATR13 promote disease susceptibility in *Arabidopsis thaliana*. *The Plant Cell* **19:** 4077-4090.

Sonnhammer EL. Durbin R. 1995. A dot-matrix program with dynamic threshold control suited for genomic DNA and protein sequence analysis. *Gene* **167:** GC1–G10.

Spanu PD, Abbott JC, Amselem J, Burgis TA, Soanes DM, Stüber K, Ver Loren van Themaat E, Brown JK, Butcher SA, Gurr SJ, Lebrun MH, Ridout CJ, Schulze-Lefert P, Talbot NJ, Ahmadinejad N, Ametz C, Barton GR, Benjdia M, Bidzinski P, Bindschedler LV, Both M,

Brewer MT, Cadle-Davidson L, Cadle-Davidson MM, Collemare J, Cramer R, Frenkel O, Godfrey D, Harriman J, Hoede C, King BC, Klages S, Kleemann J, Knoll D, Koti PS, Kreplak J, López-Ruiz FJ, Lu X, Maekawa T, Mahanil S, Micali C, Milgroom MG, Montana G, Noir S, O'Connell RJ, Oberhaensli S, Parlange F, Pedersen C, Quesneville H, Reinhardt R, Rott M, Sacristán S, Schmidt SM, Schön M, Skamnioti P, Sommer H, Stephens A, Takahara H, Thordal-Christensen H, Vigouroux M, Wessling R, Wicker T, Panstruga R. 2010. Genome expansion and gene loss in powdery mildew fungi reveal tradeoffs in extreme parasitism. *Science* **330:** 1543–1546.

Spanu PD. 2014. Messages from powdery mildew DNA: how the interplay with a host moulds pathogen genomes. *Journal of Integrative Agriculture* **13:** 233-236.

Spanu PD. 2017. Cereal immunity against powdery mildews targets RNase-Like Proteins associated with Haustoria (RALPH) effectors evolved from a common ancestral gene. *New Phytologist* **213:** 969-971.

Srichumpa P, Brunner S, Keller B, Yahiaoui N. 2005. Allelic series of four powdery mildew resistance genes at the *Pm3* locus in hexaploid bread wheat. *Plant Physiology* **139:** 885-895.

Stahl EA, Dwyer G, Mauricio R, Kreitman M, Bergelson J. 1999. Dynamics of disease resistance polymorphism at the *Rpm1* locus of *Arabidopsis*. *Nature* **400:** 667-671.

Stamatakis A. 2014. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* **30:** 1312–1313.

Staskawicz B, Dahlbeck D, Keen N, Napoli C. 1987. Molecular characterization of cloned avirulence genes from race 0 and race 1 of *Pseudomonas syringae* pv. *glycinea*. *Journal of Bacteriology* **169:** 5789-5794.

Stam R, Mantelin S, McLellan H, Thilliez G. 2014. The role of effectors in nonhost resistance to filamentous plant pathogens. *Frontiers in Plant Science* **5:** 582.

Stein M, Dittgen J, Sánchez-Rodríguez C, Hou BH, Molina A, Schulze-Lefert P, Lipka V, Somerville S. 2006. *Arabidopsis* PEN3/PDR8, an ATP binding cassette transporter, contributes to nonhost resistance to inappropriate pathogens that enter by direct penetration. *The Plant Cell* **18:** 731-746.

Steinbrenner AD, Goritschnig S, Staskawicz BJ. 2015. Recognition and Activation Domains Contribute to Allele-Specific Responses of an Arabidopsis NLR Receptor to an Oomycete Effector Protein. *PLOS Pathogens* **11:** e1004665.

Stirnweis D, Milani SD, Jordan T, Keller B, Brunner S. 2014a. Substitutions of two amino acids in the nucleotide-binding site domain of a resistance protein enhance the hypersensitive response and enlarge the PM3F resistance spectrum in wheat. *Molecular Plant-Microbe Interactions* **27:** 265-276.

Stirnweis D, Milani SD, Brunner S, Herren G, Buchmann G, Peditto D, Jordan T, Keller B. 2014b. Suppression among alleles encoding nucleotide-binding-leucine-rich repeat resistance proteins interferes with resistance in F1 hybrid and allele-pyramided wheat plants. *Plant Journal* **79:** 893-903.

Stukenbrock EH, McDonald BA. 2008. The origins of plant pathogens in agro-ecosystems. *Annual Review of Phytopathology* **46:** 75-100.

Stukenbrock EH. 2016. Hybridization speeds up the emergence and evolution of a new pathogen species. *Nature Genetics* **48:** 113-115.

Sun Q, Collins NC, Ayliffe M, Smith SM, Drake J, Pryor T, Hulbert SH. 2001. Recombination between paralogues at the *rp1* rust resistance locus in maize. *Genetics* **158:** 433-438.

Tajima F. 1989. Statistical methods to test for nucleotide mutation hypothesis by DNA polymorphism. *Genetics* **123:** 585-595.

Takken F, Rep M. 2010. The arms race between tomato and *Fusarium oxysporum*. *Molecular Plant Pathology* **11:** 309–14.

Takagi H, Abe A, Yoshida K, Kosugi S, Natsume S, Mitsuoka C, Uemura A, Utsushi H, Tamiru M, Takuno S, Innan H, Cano LM, Kamoun S, Terauchi R. 2013. QTL-seq: rapid mapping of quantitative trait loci in rice by whole genome resequencing of DNA from two bulked populations. *Plant Journal* **74:** 174–183.

Tamura K, Stecher G, Peterson D, Filipski A, Kumar S. 2013. MEGA6: Molecular Evolutionary Genetics Analysis version 6.0. *Molecular Biology and Evolution* **30:** 2725-2729.

Terauchi R, Yoshida K, Saitoh H, Kanzaki H, Okuyama Y, Fujisaki K, Miya A, Abe A, Tamiru M, Tosa Y. 2011. Studying genome-wide DNA polymorphisms to understand *Magnaporthe*–rice interactions. *Australasian Plant Pathology* **40:** 328–334.

Thomma BPHJ, van Esse HP, Crous PW, de Wit PJGM. 2005. *Cladosporium fulvum* (syn. *Passalora fulva*), a highly specialized plant pathogen as a model for functional studies on plant pathogenic Mycosphaerellaceae. *Molecular Plant Pathology* **6:** 379-393.

Thrall PH, Burdon JJ. 2000. Effect of resistance variation in a natural plant host–pathogen metapopulation on disease dynamics. *Plant Pathology* **49:** 767-773.

Thrall PH, Burdon JJ. 2003. Evolution of virulence in a plant host-pathogen metapopulation. *Science* **299:** 1735-1737.

Thrall PH, Laine AL, Ravensdale M, Nemri A, Dodds PN, Barrett LG, Burdon JJ. 2012. Rapid genetic change underpins antagonistic coevolution in a natural host-pathogen metapopulation. *Ecology Letters* **15:** 425-435.

Tori K, Cheriyan M, Pedamallu CS, Contreras MA, Perler FB. 2012. The *Thermococcus kodakaraensis Tko* CDC21-1 intein activates its N-terminal splice junction in the absence of a conserved histidine by a compensatory mechanism. *Biochemistry* **51:** 2496–2505.

Tosa Y, Tamba H, Tanaka K, Mayama S. 2006. Genetic analysis of host species specificity of *Magnaporthe oryzae* isolates from rice and wheat. *Phytopathology* **96:** 480-484.

Troch V, Audenaert K, Vanheule A, Bekaert B, Höfte M, Haesaert G. 2014. The importance of non-penetrated papillae formation in the resistance response of triticale to powdery mildew (*Blumeria graminis*). *Plant Pathology* **63:** 129-139.

Valent B, Farrall L, Chumley FG. 1991. *Magnaporthe grisea* genes for pathogenicity and virulence identified through a series of backcrosses. *Genetics* **127:** 87-101.

van den Burg HA, Harrison SJ, Joosten MH, Vervoort J, de Wit PJ. 2006. *Cladosporium fulvum Avr4* protects fungal cell walls against hydrolysis by plant chitinases accumulating during infection. *Molecular Plant Microbe Interactions* **19:** 1420–1430.

Van der Hoorn RA, De Wit PJ, Joosten MH. 2002. Balancing selection favors guarding resistance proteins. *Trends in Plant Science* **7:** 67-71.

van der Hoorn RA, Kamoun S. 2008. From guard to decoy: a new model for perception of plant pathogen effectors. *The Plant Cell* **20:** 2009-2017.

Vleeshouwers VGAA, Raffaele S, Vossen JH, Champouret N, Oliva R, Segretin ME, Rietman H, Cano LM, Lokossou A, Kessel G, Pel MA, Kamoun S. 2011. Understanding and exploiting late blight resistance in the age of effectors. *Annual Review of Phytopathology* **49:** 507–531.

Vleeshouwers VG, Oliver RP. 2014. Effectors as tools in disease resistance breeding against biotrophic, hemibiotrophic, and necrotrophic plant pathogens. *Molecular Plant Microbe Interactions* **27:** 196–206.

Wagner S, Stuttmann J, Rietz S, Guerois R, Brunstein E, Bautor J, Niefind K, Parker JE. 2013. Structural basis for signaling by exclusive EDS1 heteromeric complexes with SAG101 or PAD4 in plant innate immunity. *Cell Host & Microbe* **14:** 619-630.

Wang CIA, Guncar G, Forwood JK, Teh T, Catanzariti AM, Lawrence GJ, Loughlin FE, Mackay JP, Schirra HJ, Anderson PA, Ellis JG, Dodds PN, Kobe B. 2007. Crystal structures of flax rust avirulence proteins AvrL567-A and -D reveal details of the structural basis for flax disease resistance specificity. *The Plant Cell* **19:** 2898–2912.

Weadick CJ, Chang BS. 2012. An improved likelihood ratio test for detecting site-specific functional divergence among clades of protein-coding genes. *Molecular Biology and Evolution* **29:** 1297-1300.

Whisson SC, Boevink PC, Moleleki L, Avrova AO, Morales JG, Gilroy EM, Armstrong MR, Grouffaud S, Van West P, Chapman S, Hein I. 2007. A translocation signal for delivery of oomycete effector proteins into host plant cells. *Nature* **450:** 115-8.

Wolfe MS, Schwarzbach E. 1978. Patterns of race changes in powdery mildews. *Annual Review of Phytopathology* **16:** 159-180.

Wicker T, Oberhaensli S, Parlange F, Buchmann JP, Shatalina M, Roffler S, Ben-David R, Doležel J, Šimková H, Schulze-Lefert P, Spanu PD, Bruggmann R, Amselem J, Quesneville H, Ver Loren van Themaat E, Paape T, Shimizu KK, Keller B. 2013. The wheat powdery mildew genome shows the unique evolution of an obligate biotroph. *Nature Genetics* **45:** 1092-1096.

Wiermer M, Feys BJ, Parker JE. 2005. Plant immunity: the EDS1 regulatory node. *Current Opinion in Plant Biology* **8:** 383-389.

Wilcoxon F. 1945. Individual comparisons by ranking methods. *Biometrics Bulletin* **1:** 80-83.

Williams SJ, Sornaraj P, deCourcy-Ireland E, Menz RI, Kobe B, Ellis JG, Dodds PN, Anderson PA. 2011. An autoactive mutant of the M flax rust resistance protein has a preference for binding ATP, whereas wild-type M protein binds ADP. *Molecular Plant Microbe Interactions* **24:** 897-906.

Wu Y, Close TJ, Lonardi S. 2008. On the accurate construction of consensus genetic maps. *Computational Systems Bioinformatics Conference* **7:** 285-296.

Xiao S, Wang W, Yang X. 2008. Evolution of resistance genes in plants. In: Holger, eds. *Innate immunity of plants, animals, and humans*. Heidelberg, Germany: Springer Berlin, 1-25.

Yaegashi H, Udagawa S. 1978. The taxonomical identity of the perfect state of *Pyricularia grisea* and its allies. *Canadian Journal of Botany* **56:** 180-183.

Yahiaoui N, Srichumpa P, Dudler R, Keller B. 2004. Genome analysis at different ploidy levels allows cloning of the powdery mildew resistance gene *Pm3b* from hexaploid wheat. *The Plant Journal* **37:** 528-538.

Yahiaoui N, Brunner S, Keller B. 2006. Rapid generation of new powdery mildew resistance genes after wheat domestication. *The Plant Journal* **47:** 85–98.

Yahiaoui N, Kaur N, Keller B. 2009. Independent evolution of functional *Pm3* resistance genes in wild tetraploid wheat and domesticated bread wheat. *The Plant Journal* **57:** 846–856.

Yang Z, Nielsen R. 2000. Estimating synonymous and nonsynonymous substitution rates under realistic evolutionary models. *Molecular Biology and Evolution* **17:** 32-43.

Yang Z, Wong WS, Nielsen R. 2005. Bayes empirical Bayes inference of amino acid sites under positive selection. *Molecular Biology and Evolution* **22:** 1107-1118.

Yang Z. 2007. PAML 4: phylogenetic analysis by maximum likelihood. *Molecular Biology and Evolution* **24:** 1586-1591.

Yoshida K, Saitoh H, Fujisawa S, Kanzaki H, Matsumura H, Yoshida K, Tosa Y, Chuma I, Takano Y, Win J, Kamoun S, Terauchi R. 2009. Association genetics reveals three novel avirulence genes from the rice blast fungal pathogen *Magnaporthe oryzae*. *The Plant Cell* **21:** 1573-1591.

Yoshida K, Saunders DG, Mitsuoka C, Natsume S, Kosugi S, Saitoh H, Inoue Y, Chuma I, Tosa Y, Cano LM, Kamoun S, Terauchi R. 2016. Host specialization of the blast fungus *Magnaporthe oryzae* is associated with dynamic gain and loss of genes linked to transposable elements. *BMC Genomics* **17:** 370.

Yun BW, Atkinson HA, Gaborit C, Greenland A, Read ND, Pallas JA, Loake GJ. 2003. Loss of actin cytoskeletal function and EDS1 activity, in combination, severely compromises non-host resistance in *Arabidopsis* against wheat powdery mildew. *The Plant Journal* **34:** 768-777.

Zeng FS, Yang LJ, Gong SJ, Shi WQ, Zhang XJ, Wang H, Xiang LB, Xue MF, Yu DZ. 2014. Virulence and Diversity of *Blumeria graminis* f. sp. *tritici* Populations in China. *Journal of Integrative Agriculture* **13:** 2424-2437.

Zhang Z, Feechan A, Pedersen C, Newman MA, Qiu JL, Olesen KL, Thordal-Christensen H. 2007. A SNARE-protein has opposing functions in penetration resistance and defence signalling pathways. *The Plant Journal* **49:** 302-312.

Zhang Z, Lenk A, Andersson MX, Gjetting T, Pedersen C, Nielsen ME, Newman MA, Hou BH, Somerville SC, Thordal-Christensen H. 2008. A lesion-mimic syntaxin double mutant in *Arabidopsis* reveals novel complexity of pathogen defense signaling. *Molecular Plant* **1:** 510-527.

Zhang NW, Lindhout P, Niks RE, Jeuken MJW. 2009. Genetic dissection of *Lactuca saligna* nonhost resistance to downy mildew at various lettuce developmental stages. *Plant Pathology* **58:** 923-932.

Zhang WJ, Pedersen C, Kwaaitaal M, Gregersen PL, Mørch SM, Hanisch S, Kristensen A, Fuglsang AT, Collinge DB, Thordal-Christensen H. 2012. Interaction of barley powdery mildew effector candidate CSEP0055 with the defence protein PR17c. *Molecular Plant Pathology* **13:** 1110-1119.

Zipfel C. 2014. Plant pattern-recognition receptors. *Trends in Immunology* **35:** 345-351.

Acknowledgements

In the beginning there was Francis and Roi. Their enthusiasm and interest in our work convinced me to come to Switzerland, and I have them to thank that I ended up on this project. Although Roi left only a month after I arrived, we have continued to be friends and collaborators. I enjoyed your laughter, passion for pastries, and even your tubes, boxes, and excel files written and labeled in Hebrew. דותה!

Soon after I arrived I met my quiet, grumpy, exceedingly lovely friend Alexey. Thanks to him for all the dry sarcasm that got me through the early days. I sorely miss all the long rants you listened to while we ice skated, but I still have reminders of your advice all over my desk, including my favorite prancing kitten with a healthy attitude.

Justine arrived not long after me. I have her to thank for a lot of good times over the years: hikes, parties, trips to Paris, team running events… And especially the last time we sat in Franzos having a glass of wine. I will forever think of you when I'm out hiking and I see a marmot!

Coraline is my PhD-sister. We supported each other through the early days of mapping, and from there grew into our now very different projects. We've bonded over all of the joys and pains of working with Salim. Thanks for all of the emotional, technical, logistical, and culinary support over the years! I hear your voice every time I buy tomatoes out of season. (Forgive me!)

Both Fabrizio and Javi arrived on the same day, and ever since they have been a constant moral and intellectual support. It has been a joy to grow close to you both. Thanks for sharing your experience with me and for being an attentive ear when I needed to share mine. Write me whenever you want to get a beer- I will be writing you!

Anja and I found each other the way kindred spirits do. Despite the monumental divide between P2 and P3 we came together as good friends and confidants. I always appreciated her outsiders view on things in stressful lab situations. I love discussing our big plans over coffee after yoga and I'm looking forward to sharing our next adventures in life together, too!

Linda, dear Linda. Not only was she an amazing and entertaining friend (how big is your casserole dish, again?), but she was a huge help for a lot of the work in this thesis. It's because of her that I can even defend in 2017. *Thank you so much!*

Only later came the full Aquarium crew. Stefan, Luisa, Marion, Lukas… And most recently Andreas and Seraina. You are fantastic colleagues, friends, drinking buddies, co-conspirators… I couldn't think of a better group to arrive to (sometimes early) and leave from (sometimes quite late) every day. Please continue inviting me to game night and after work drinks! I will miss you!

Salim- You didn't think I left you out, did you? It has been a strange 5 years getting to know and work with you. It's been a roller-coaster for at least the last 4 of those years. Although you have not been the easiest friend to work with, you have, most of the time, when you were trying, and in a good mood- not too tired, not too stressed, not too hungry/angry/busy, been a good friend. Sometimes too good. You know, the kind of 'suffering is good for you' kind of good. Or the 'I expect more because I think you can do more' kind of good. We both learned a lot from each other, grew incredibly *incredibly* close and all in all, in the end, things are good. So thanks for everything, it definitely hasn't been boring. I will miss how funny you are when you're stressed, and your quite dark humor. Thanks for never minding that I'm not very funny. Just don't tell anyone about my Grandpa's horse, ok? And I won't say anything about toes.

I really truly would not have been here or done any of this (really rather too large) work without Beat. He flew me in from far away, paid my way through Masters and PhD work, and was a great support in this last phase of writing up the thesis. It wasn't always easy working together, but I learned a lot from that and I think we are both quite happy about what has come out and where the project is headed. I am very happy to have made what I hope is a meaningful contribution to the work in your group.

Thanks a lot to all the other people I haven't mentioned but whose smiles, '*Morge!*'s, secretarial skills, technical assistance, gardening expertise, menus, beers, projects, experiences, ideas, and general camaraderie has made the past 5 years so enjoyable.

I owe many thanks to my family, who so generously (though not without complaint) sacrificed some of their happiness to allow me to live abroad, and to visit much too rarely because of work. I know it is because you have always believed in me that I can believe in myself.

Finally, I would like to express my appreciation for the continuous patience, understanding, love, and support that Mark has given me over these past 5 years. When I thought I complained too much, he listened more. When I thought my low moods would bring him down, he tried harder to cheer me up. I will never be able to express enough my gratitude for the strength of your encouragement and support over the years- I just hope I will be able to return it someday.

Curriculum vitae

Surname:	MC NALLY
Name:	Kaitlin Elyse
Date of birth:	23.10.1987
Plant of origin:	Massillon, Ohio, United States of America

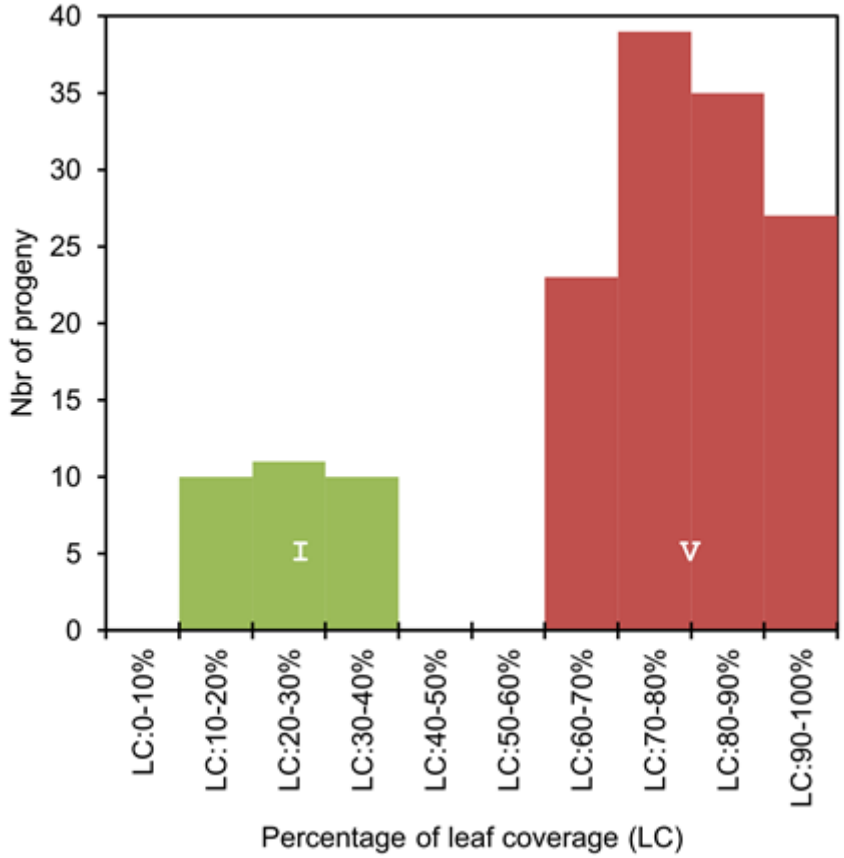
Education

2006-2010	Bachelor of Science in Genetics Certificate in International Agriculture University of Georgia Athens, Georgia, USA
2012-2013	Master of Science in Plant Science Under the supervision of Prof. Beat Keller Institute of Plant and Microbial Biology, University of Zürich Title: Molecular analysis of the AvrPm3-Pm3 interaction in the wheat powdery mildew pathosystem: mapping and identification of *AvrPm3f2* candidate avirulence gene
2013-2017	PhD thesis, Science and Policy Program Under the supervision of Prof. Beat Keller Institute of Plant and Microbial Biology, University of Zürich Title: Molecular and Evolutionary Studies of Race-specific Avirulence Factors from Wheat Powdery Mildew (*Blumeria graminis* f. sp. *tritici*)

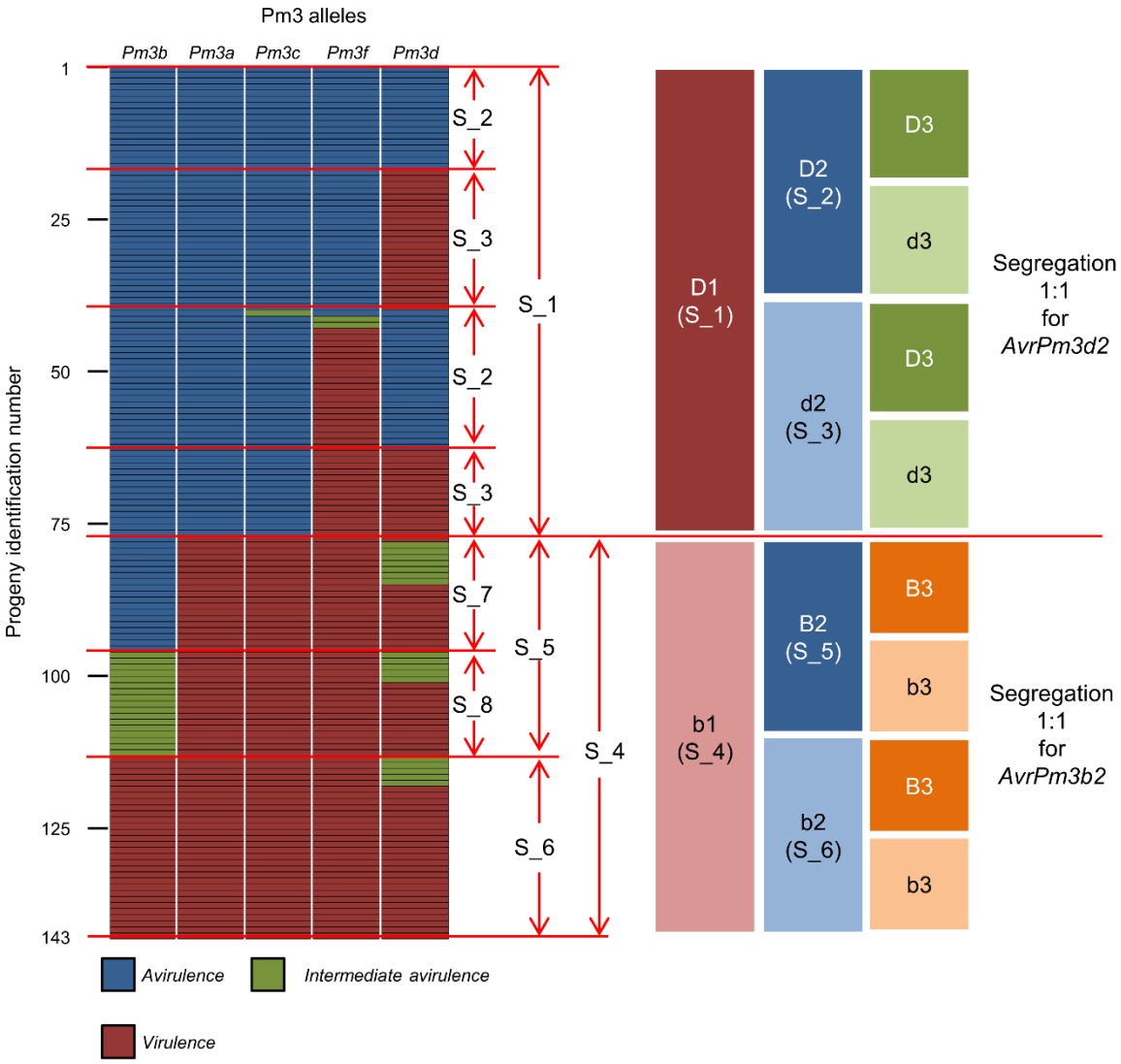
Employed as a PhD student in the FastTrack program at UZH since 22. June 2012.

Appendix I – Supplemental data for Chapter 3

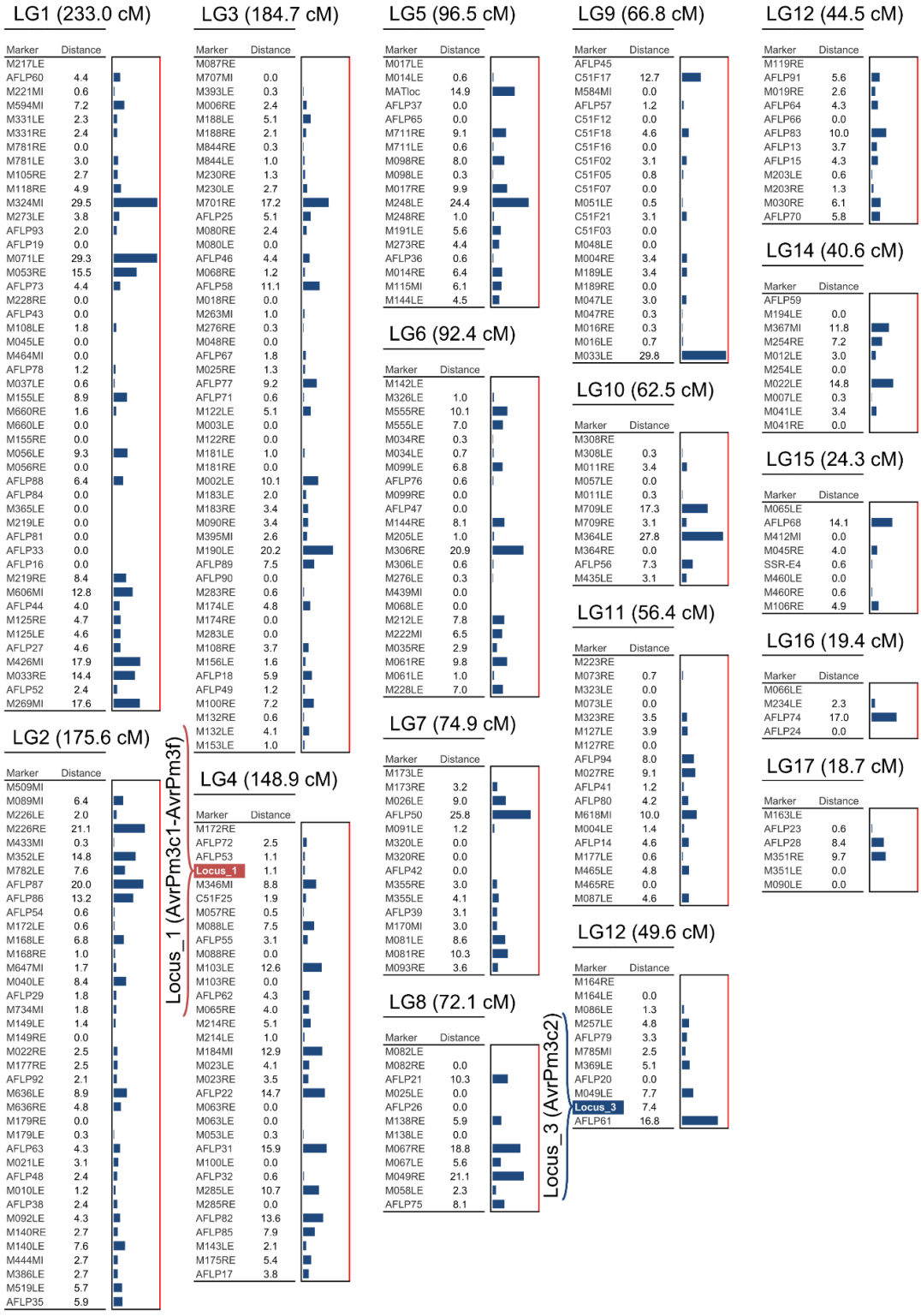
Supplemental Figures for Chapter 3



Supplemental Figure 1. Assignment of powdery mildew phenotypes based on the percentage of leaf coverage. Leaf coverage (LC) scores corresponding to intermediate avirulence (I) and virulence (V) on *Pm3b* and *Pm3d* are represented. LCs grouped into three distinct classes where progeny with LC=0% (not represented) were designated as avirulent, LC=10-40% designated as intermediate avirulent and LC=60-100% as virulent.



Supplemental Figure 2. Selection of progeny subsets for mapping *AvrPm3b2* and *AvrPm3d2*. On the left, progeny phenotypes are color coded as indicated and progeny subsets are indicated by 'S' followed by an identification number. On the right, progeny genotypes are indicated as well as the subset of the population harboring this genotype (see Supplemental Text). Progeny subsets segregating in a 1:1 ratio for $AvrPm3^{b2}$ and $AvrPm3^{d2}$ are indicated. The colors and codes used for *Avr* and *avr* loci are the same as in Figure 2A.



Supplemental Figure 3. SNP and AFLP marker based genetic map of the 96224 x JIW2 mapping population. The 96224 x JIW2 genetic linkage map consists of 338 markers mapping **(Supplemental Figure 3. Continued from previous page)** in 17 linkage groups. Linkage was considered significant at a threshold of 30 cM, as indicated by the red line. For simplicity, the loci are labeled and identified by the genotype of the avirulent parent. *Locus_1* (*AvrPm3*^{c1}-*AvrPm3f*) maps to linkage group 4 and *locus_3* (*AvrPm3*^{c2}) to linkage group 12. The location of the mating-type locus is given (MATloc, LG5). KASP markers are labeled 'M' for 'marker,' followed by the FPC contig number, and 'LE' (left end), 'RE' (right end), or 'MI' (middle), specifying the relative location of the marker within the contig. AFLP markers are also labeled, and the remaining markers were derived from different mapping projects. The distance between each marker is given in cM and illustrated by a horizontal blue bar.

Supplemental Tables for Chapter 3

Supplemental Table 1. Primers used for gene annotation, genotyping and molecular cloning.

Primers ID	5' to 3' Sequence
pu7_CAPS_F1	CAATCTAGCAATACGCTCTGGACATAG
pu7_CAPS_R1	GGAAAGGACAGAAATTTGGTTCAAG
pu7_seq_F2	CCTCTGAACCGCCCCATTT
pu7_seq_R2	CAAATCAGGTCACCCCACCA
pu7_RACE_F1	GAGTGGCCCTGTCGCTAACGCTTC
pu7_RACE_F2	GCACTTCACCTAACGGACTAAGGAAAGG
pu7_RACE_R1	GTTTAATTGAGGCCTCTCTGTTGTTGG
pu7_RACE_R2	GCAATACGCTCTGGACATAAGCCTGTAG
pu7_TOPOwSP_F	CACCATGAGAACCTTCAGTCTTG
pu7_TOPOnoSP_F	CACCATGGGCCCTGTCGCTAAC
pu7_TOPO_R	CTAGTGCAGAATTATGTTTAATTGAGG
pu7_G84E_F	ATCTTAAAAAGAGAAGAAGAAAATA
pu7_G84E_R	TACATTAAAGTATCTTGAACCAAAT
pu7_A107V_F	GAAATTATTTATCTACAGGTTTATGTCC
pu7_A107V_R	CTTCAAATAGTTGATTCCCACA
Bcg1gdnaF1	TTCATCTTCCCATCCACGTC
Bcg1gdnaR1	AGGGCCAGTTTGTTCATGTC
Bcg1gdnaF2	TCTTCATCTTCCCATCCACGTC
Bcg1gdnaR2	CGTACAGGTAGTTGAGGGCCAGT

The 'CAPS' primers were designed to amplify a sequence within *Pu_7* that is cleaved by HaeIII in PCR amplicons from the 96224 genotype, but not the 94202 one. The 'seq' primers are located in the 5' and 3' UTRs and were designed to verify the complete sequence of *Pu_7*. The 'RACE' primers are nested primers used to confirm the cDNA sequence of the expressed *Pu_7* using to the SMARTer RACE cDNA kit (Takara Bio, Shiga, Japan) protocol. The 'TOPO' primers were used to clone *Pu_7* with (wSP) and without (noSP) signal peptide into the pENTR vector using the pENTR/D-TOPO Cloning Kit (Life Technologies, Carlsbad, USA) according to the manufacturer. The 'G84E' and 'A107V' primers were used to perform site directed mutagenesis on *Pu_7*. The Bcg1gdna primers were used for amplification and sequencing of specific *Bcg1* amplicons from progeny genomic DNA.

Supplemental Table 2. BAC contigs associated to avirulence of the 96224 isolate on *Pm3f*.

FPC-contigs[1]	Linkage[2]	Marker ID[3]	Genetic distance[4]
Ctg-172	*AvrPm3f1*	M172RE	4.7 cM
Ctg-346	*AvrPm3f1*	M346MI	15.3 cM
Ctg-057	*AvrPm3f1*	M057RE	17.3 cM
Ctg-413	*AvrPm3f1*	*K413_2*	25.4 cM
Ctg-125	*AvrPm3f2*	M125RE	14.9 cM
		M125LE	10 cM
Ctg-026	*AvrPm3f2*	*K26_04*	6.2 cM
		K26_03	6.2 cM
		K26_06	6.2 cM
		K26_02	6.2 cM
Ctg-426	*AvrPm3f2*	*K426_3*	6.2 cM
		K426_4	6.2 cM
		K426_6	4.6 cM
		M426MI	4.6 cM
Ctg-052	*AvrPm3f2*	*K52_02*	1.5 cM
		K52_24	0.0 cM
		K52_09	0.0 cM
		K52_05	0.0 cM
		K52_10	0.0 cM
		K52_07	0.0 cM
Ctg-033	*AvrPm3f2*	*K33_11*	0.8 cM
		M033RE	12.3 cM

[1] BAC-contigs identified by BSA as associated with avirulence towards *Pm3f*. Genetic association was determined based on the percentage of SNP positions corresponding to the genotype of the avirulent parent 96224 (see Methods).
[2] Genetic linkage to *AvrPm3f1* (*locus_1*) or *AvrPm3f2* (*locus_2*) was determined based on contig-specific genetic markers indicated in the third column.
[3] Contig specific genetic markers. KASP markers are indicated by 'M' followed by contig-number and relative physical position of the marker indicated as 'LE' (Left end), 'RE' (Right end) or 'MI' (middle). CAPS markers are italicized and indicated by 'K' followed by contig-number and marker identifier.
[4] Genetic distances from *AvrPm3f1* or *AvrPm3f2* of contig-specific markers indicated in centiMorgans (cM).

Supplemental Table 3. RT-qPCR primers.

Genes	Primers ID	5' to 3' sequence
Reference genes:		
Act1 (*Actin*)	qRT-BgtAct1F	GACAATGGGTGGTCGGGAAT
	qRT-BgtAct1R	CAACACCATGTTCGATGGGATA
Gapdh (*Glyceraldehyde 3-phosphate dehydrogenase*)	qRT-BgtGapdhF	TGTCTTCCGAAACGCTGCTC
	qRT-BgtGapdhR	AGTCCGTCCTCGACTGCTTGT
Target genes:		
Bcg1	qRT-BgtBcg1F	TTCCTGGAGTAGTTTTCTACCA
	qRT-BgtBcg1R	TGTAAGCCGTTGGGTAAAG
Pu_7	qRT-BgtPu_7F	GCACTTCACCTAACGGACTAAGGAAAGG
	qRT-BgtPu_7R	GCAATACGCTCTGGACATAAGCCTGTAG

Supplemental Text for Chapter 3

Mapping the AvrPm3^{c2} locus involved in the Pm3c interaction in the JIW2 population

In the JIW2 population studied by Parlange et al. (2015) the F_1 segregation on *Pm3c* fitted a 2:1:1 (A:I:V) ratio, which is consistent with a model where two loci differing between the parents are involved in the interaction. Parlange et al. (2015) have also demonstrated that one of these loci (*AvrPm3^{c1}*) was identical to the locus controlling the *AvrPm3f-Pm3f* interaction in the JIW2 population. This locus was mapped in the JIW2 population (Parlange et al., 2015) and corresponds to *locus_1* in the 94202 population described here. Parlange et al. (2015) also predicted that the intermediate phenotype on *Pm3c* was controlled by *AvrPm3^{c2}*, the second locus interacting with *Pm3c*, meaning that intermediate and virulent progeny on *Pm3c* segregate in 1:1 ratio for *AvrPm3^{c2}*. Therefore, to map the locus corresponding to *AvrPm3^{c2}* in the JIW2 population, we used the subset of progeny that were intermediate or virulent on *Pm3c*. *AvrPm3^{c2}* mapped to the genetic region defined by markers M049LE and AFLP61 and is referred to as *locus_3* in this study (see Supplemental Figure 5).

Mapping the locus controlling the AvrPm3^{b2}-Pm3b and AvrPm3^{d2}-Pm3d interactions in the 94202 population and phenotype to genotype assignment

Phenotypic segregation on *Pm3d* followed a 2:1:5 (A:I:V) ratio, indicating that three independently segregating loci are involved in the *AvrPm3d-Pm3d* interaction. There are 8 possible genotypes for three loci, each existing in two allelic *Avr* and *avr* forms (Figure 1). According to the 2:1:5 ratio, two of eight genotypes are avirulent, thus avirulence can only result from a combination of two *Avr* loci. One of these two loci must be the avirulent *locus_1*, as concluded from the observations that all avirulent progeny on *Pm3d* are also avirulent on *Pm3a* and *Pm3c* (the phenotype for these alleles only depends on the genotype at *locus_1*, which was mapped for *Pm3a*, *Pm3c* and also *Pm3e* based on the 1:1 segregation ratio on these alleles, Figure 1A). This was confirmed for *Pm3d* by mapping *AvrPm3^{d1}* using the S_1 and S_4 subsets, where we know that the S_3 subset has the genotype *AvrPm3^{d1}*, thus obtaining a population segregating in a 1:1 ratio for *AvrPm3^{d1}* (see Supplemental Figure 2). The S_1 progeny, which all possess the avirulent genotype at *locus_1* (*AvrPm3^{d1}*), segregate in 1:1 (A:V) ratio on *Pm3d*. Since two *Avr* loci are required for avirulence on *Pm3d*, this ratio can only be explained by a 1:1 segregation for *AvrPm3^{d2}*, the second locus required for avirulence (see Supplemental Figure 2). We used the S_1 progeny to map *AvrPm3^{d2}*, which located to the same genetic region as *AvrPm3^{c2}* at *locus_3* (see Supplemental Figure 5). We were also able to assign which genotype combination is responsible for intermediate avirulence on *Pm3d*. Phenotype data indicate that all the progeny with an intermediate avirulence phenotype on *Pm3d* belong to the S_4 subset, which are all virulent on *Pm3a* and *Pm3c*, thus possessing the virulent genotype (*avrPm3^{d1}*) at *locus_1* (see Supplemental Figure 2). Here, the 2:1:5 phenotypic segregation on *Pm3d* is the key observation, since it shows that none of the three independently segregating loci is individually capable of conferring recognition by *Pm3d*, which results either in an avirulent or intermediate avirulent phenotype. Thus, in Figure 1, we assigned intermediate avirulence to the only genotype combination involving two *Avr* loci in the S_4 progeny: *avrPm3^{d1}-AvrPm3^{d2}-AvrPm3^{d3}* (indicated d1-D2-D3 in Figure 1).

Phenotypic segregation on *Pm3b* followed a 5:1:2 (A:I:V) ratio, indicating that three independently segregating loci are involved in the *AvrPm3b-Pm3b* interaction. There are 8 possible genotypes for three loci, each existing in two allelic *Avr* and *avr* forms (Figure 1). According to the 5:1:2 ratio, five combinations result in avirulence on *Pm3b*. Four of these combinations involve the avirulent *locus_1*, since all avirulent progeny on *Pm3a* and *Pm3c* are also avirulent on *Pm3b*. This was confirmed by mapping *AvrPm3^{b1}* using the S_1 and S_4 subsets, where we know that the S_5 subset has the

genotype *avrPm3^{b1}* (see Supplemental Figure 2). Consistent with the 5:1:2 segregation pattern, an additional locus is responsible for conferring avirulence and/or intermediate avirulence towards *Pm3b*. This hypothesis was confirmed in the S_5 progeny, which possess the virulence allele at *locus_1* (*avrPm3^{b1}*) and are either avirulent (S_7) or intermediate avirulent (S8) on *Pm3b* (see Supplemental Figure 2). Here, the key observation is that the S_4 progeny segregate in a 1:1:2 (A:I:V) ratio on *Pm3b*, which indicates that two independently segregating loci, different from *locus_1*, are involved in conferring avirulence and intermediate avirulence on *Pm3b* specifically in these progeny. From this ratio we inferred that the S_4 progeny segregates in a 1:1 ratio for *AvrPm3^{b2}*, where avirulence results from the combination of *AvrPm3^{b2}* and *AvrPm3^{b3}* (B2-B3), while intermediate avirulence results from a combination of *AvrPm3^{b2}* and *avrPm3^{b3}* (B2-b3) (see Supplemental Figure 2). We used the S_4 progeny to map *AvrPm3^{b2}* which happened to map to the same genetic region as *AvrPm3^{d2}* and *AvrPm3^{c2}* at *locus_3* (see Supplemental Figure 5). Thus, in Figure 1, avirulence was assigned to the *avrPm3^{b1}-AvrPm3^{b2}-AvrPm3^{b3}* genotype (indicated b1-B2-B3) and intermediate avirulence to the *avrPm3^{b1}-AvrPm3^{b2}-avrPm3^{b3}* genotype (indicted b1-B2-b3).

High molecular weight fungal DNA extraction

To increase DNA yield, cell disruption was improved by combining thorough mechanical grinding and heat-shock as follows: 100 mg of conidia were flash frozen in liquid nitrogen and grinded in a 2ml tube using 3mm stainless-steel beads and a high speed plate shaker. After three rounds of 30 seconds grinding at a frequency of 30/s, 300 µl of a 5% Sarcosyl solution preheated at 65°C were added, followed by vigorous vortexing, then immediately supplemented with 700 µl of a solution containing 0.2 M Tris(hydroxymethyl)aminomethane at pH 7.5, 50 mM EDTA, 2 M NaCl, 2% Cetyl trimethylammonium bromide (CTAB) and 0.25 M Sodium metabisulfite ($Na_2S_2O_5$).

To improve DNA quality, samples were handled as follows: after cell disruption, tubes were vortexed vigorously and incubated 15-30 minutes at 65°C. After incubation, 600 µl of Chloroform were added to the tubes, followed by a 14,000 RCF centrifugation for 10 min at 4°C. The supernatant was collected carefully and supplemented with 1 volume of 100% Isopropanol at -20°C. Tubes were mixed by gentle inversion 6-8 times then centrifuged 10 min at 14'000 rcf, 4°C. After removing the supernatant, the pellet was dried on ice for 15-30 minutes, and then re-suspended in 450 µl of standard TrisEDTA buffer. The dissolved pellet was purified using Amicon Ultra 0.5 ml centrifugal filters MWCO 30kDa (Sigma-Aldrich, Steinheim, Germany) according to the manufacturer.

Genotyping of Bcg1 and Pu_7 in the F_1 progeny from the 94202 population.

To determine the genotype of *Pu_7* in the 94202 population, we used the pu7_CAPS marker which was designed on the gene and derived from a SNP between the parental sequences (Supplemental Table 1). This marker was used in high-resolution mapping to demonstrate co-segregation of *Pu_7* and *AvrPm3^{f2}* in the S1 subset in the 94202 population. For *Bcg1*, the presence of two paralogous genes, *Bcg6* and *Bcg7*, sharing high sequence homology with *Bcg1* (Parlange et al., submitted), did not allow to design a marker based on polymorphisms inside the gene sequence. To overcome this problem, we used the genotype data from the flanking markers F3 and F5, which define a genetic interval of 2.5 cM containing *Bcg1* (Parlange et al., Supplementary Table S2). We have also used sequence (SNP) information obtained from sequencing of *Bcg1* amplicons. To ensure specific amplification, forward and reverse primers were designed in unique sequences upstream and downstream of *Bcg1,* respectively (Supplemental Table 1). The genotype of individual progeny was assigned as follows: i) a progeny was assigned the *Bcg1^{avr}* genotype if both *Bcg1* flanking markers had the genotype of the 96224 avirulent parent, ii) a progeny was assigned the virulent *Bcg1^{vir}* genotype if

both flanking markers had the genotype of the 94202 virulent parent, iii) in the 3 progeny isolates showing recombination between the flanking markers, the genotype of *Bcg1* was resolved based on specific PCR amplification and sequencing of the gene, iv) progeny for which we had incomplete flanking marker data, the genotype of *Bcg1* was also resolved based on specific PCR amplification and sequencing of the gene.

Suppression of the AvrPm3$^{a2/f2}$-Pm3a/f interaction by Bcg1^{vir} in N. benthamiana and N. tabacum

We have performed suppression experiments using agroinfiltration assays in *N. benthamiana* (9 leaves) and *N. tabacum* (30 leaves). *AvrPm3$^{a2/f2}$*and *Pm3f$^{L456P/Y458H}$* were co-expressed together with the *Bcg1^{vir}* allele originating from the virulent parent, or the GUS reporter gene, in a three-component system. In case of suppression, we expect to see HR in the combination with GUS, serving as a control, and HR suppression in the combination including *Bcg1^{vir}*. We used an infiltration ratio of 3:1:9 (*AvrPm3$^{a2/f2}$*: *Pm3f$^{L456P/Y458H}$*: *Bcg1^{vir}*) and assessed HR 5 days after agroinfiltration. No HR was observed with both combinations for 10 leaves (26%). This is probably caused by the fact that the R:Avr ratio in the final mix is diluted by the presence of large amounts of a third construct, which is (based on our experience with this system) not optimal for developing HR. For 20 leaves (51%) we observed HR, independent of the presence or absence of the suppressor. Finally and most importantly, in the remaining 9 leaves (23%), HR suppression was observed when *Bcg1^{vir}* was present. The relevance of this result is provided by the fact that we observed no leaves with HR suppression when GUS was delivered instead of *Bcg1^{vir}*. To summarize, HR suppression was only observed when co-expressing *AvrPm3$^{a2/f2}$*, *Pm3f* $^{L456P/Y458H}$ and *Bcg1^{vir}*, which supports the hypothesis that *Bcg1^{vir}* acts as a suppressor of the *AvrPm3$^{a2/f2}$-Pm3a/f* interaction.

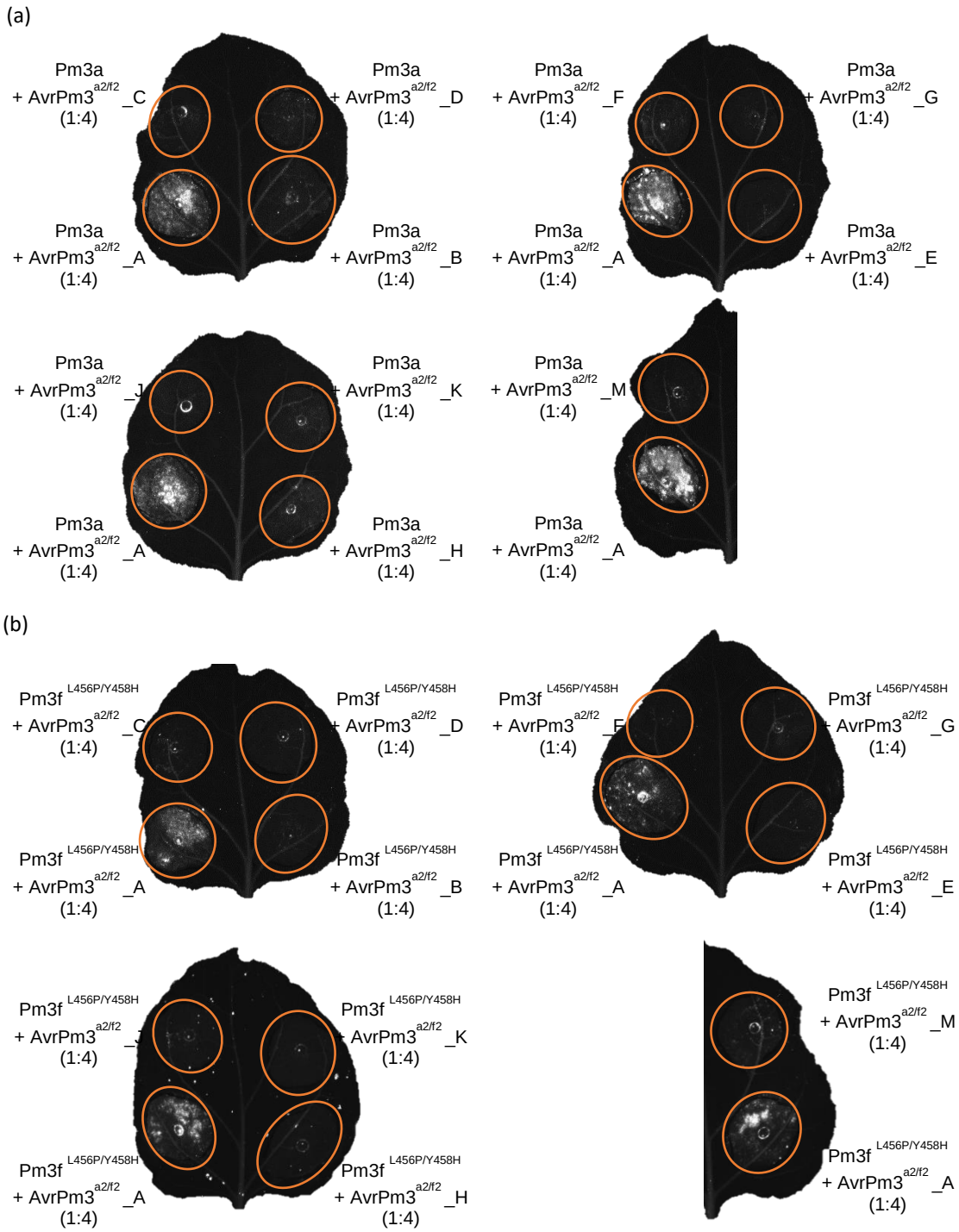


Figure S1. Recognition tests for AVRPM3$^{A2/F2}$ variants. Transient agrobacterium infiltration assays in *Nicotiana benthamiana* with a 1:4 ratio of R:AVR protein show no interaction between AVRPM3$^{A2/F2}$ variants other than 'A' (positive control) and the PM3A and PM3F$^{L456P/Y458H}$ resistance proteins from wheat. Fluorescence imaging of the leaves taken at 5 dpi reveals HR only with the AVRPM3$^{A2/F2}$ variant 'A' + PM3A or PM3F$^{L456P/Y458H}$ controls.

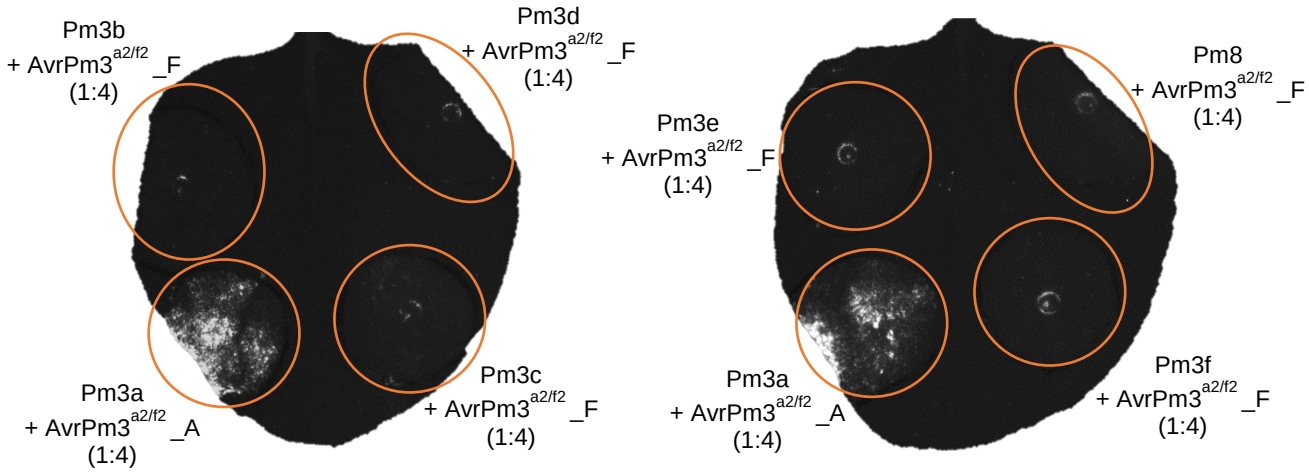


Figure S2. Recognition tests for AVRPM3$^{A2/F2}$-F. Transient agrobacterium infiltration assays in *Nicotiana benthamiana* show no interaction between the AVRPM3$^{A2/F2}$ variant 'F' and the PM3A, B, C, D, E, F$^{L456P/Y458H}$, and PM8 resistance proteins from wheat. Fluorescence imaging of the leaves taken at 5 dpi reveals HR only with the AVRPM3$^{A2/F2}$ variant 'A' + PM3A control.

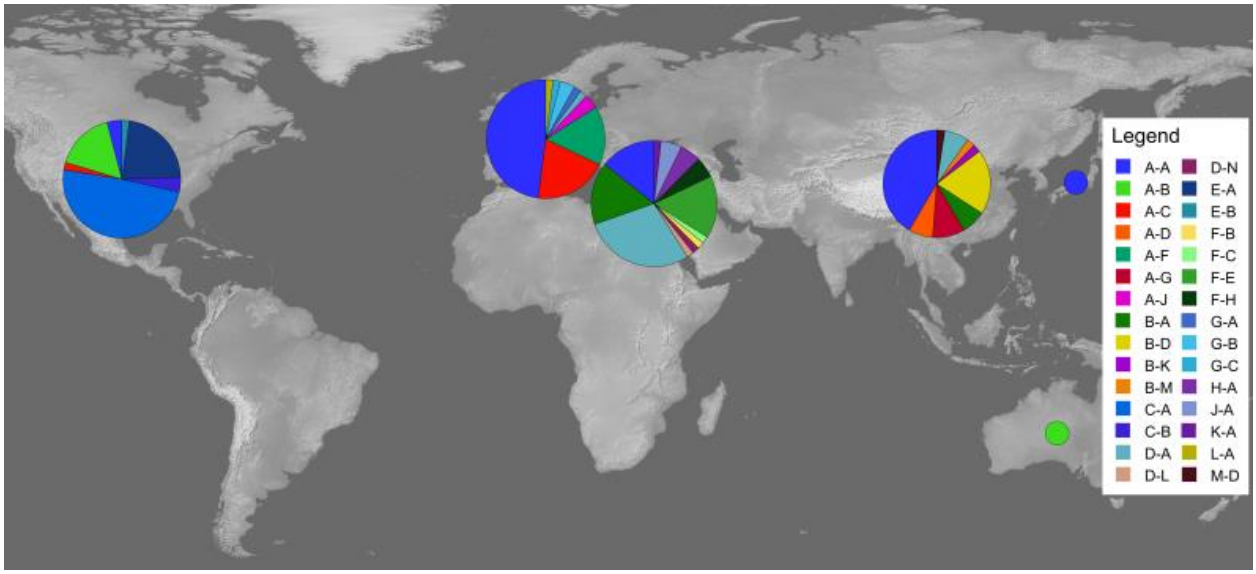


Figure S3. Geographic distribution of AVRPM3$^{A2/F2}$-SVRPM3$^{A1/F1}$ variant combinations. Variant combinations (AVR-SVR) from the US, Europe, Israel, China, Japan, and Australia. Area of the circles correspond to the number of isolates from that region.

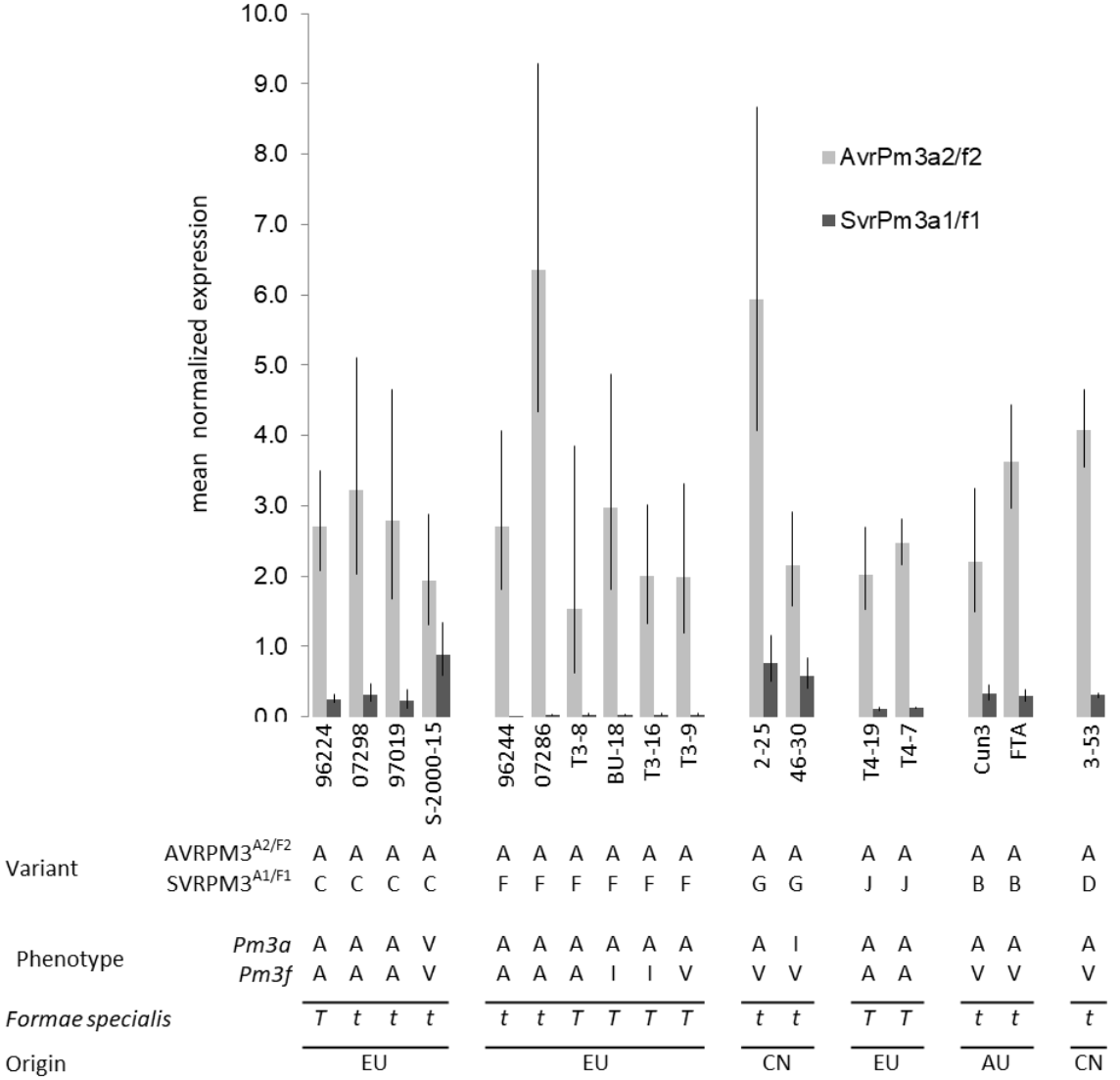


Figure S4. Expression analysis of isolates encoding the active AVRPM3$^{A2/F2}$-A variant and a SVRPM3$^{A1/F1}$ variant with unknown activity. The mean normalized expression of *SvrPm3$^{a1/f1}$* (dark grey histograms) and *AvrPm3$^{a2/f2}$* (light grey histograms) at 2d after infection. Leaf segments from the susceptible recurrent parent line '*Chancellor*' were infected with *B.g. tritici (t)* and *B.g. triticale (T)* isolates from diverse geographic origins (Europe, EU; China, CN; Australia, AU) that contain the active AVRPM3$^{A2/F2}$ and either the SVRPM3$^{A1/F1}$ -C, -F, -G, -J, -B, or -D variant.

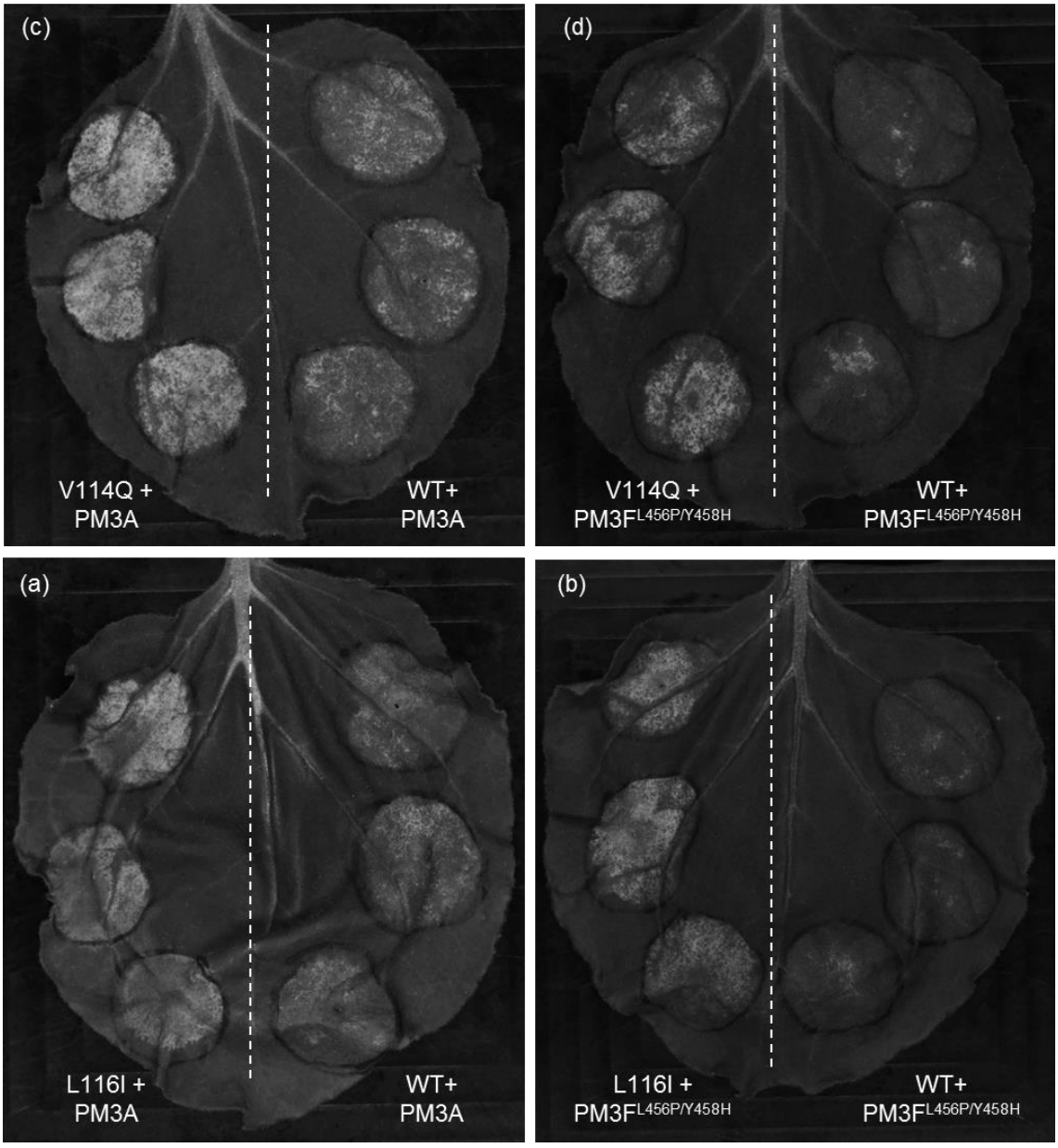


Figure S5. Single site mutations in AVRPM3$^{A2/F2}$ that enhance recognition when co-infiltrated with PM3A and PM3F$^{L456P/Y458H}$. Example images of HR quantification assays taken at 2 dpi. *Agrobacteria* expressing the V114Q and L116I altered AVRPM3$^{A2/F2}$ constructs were co-infiltrated in *Nicotiana benthamiana* with PM3A and PM3F$^{L456P/Y458H}$ and these were compared against wildtype AVRPM3$^{A2/F2}$ (WT) with PM3A and PM3F$^{L456P/Y458H}$, respectively.

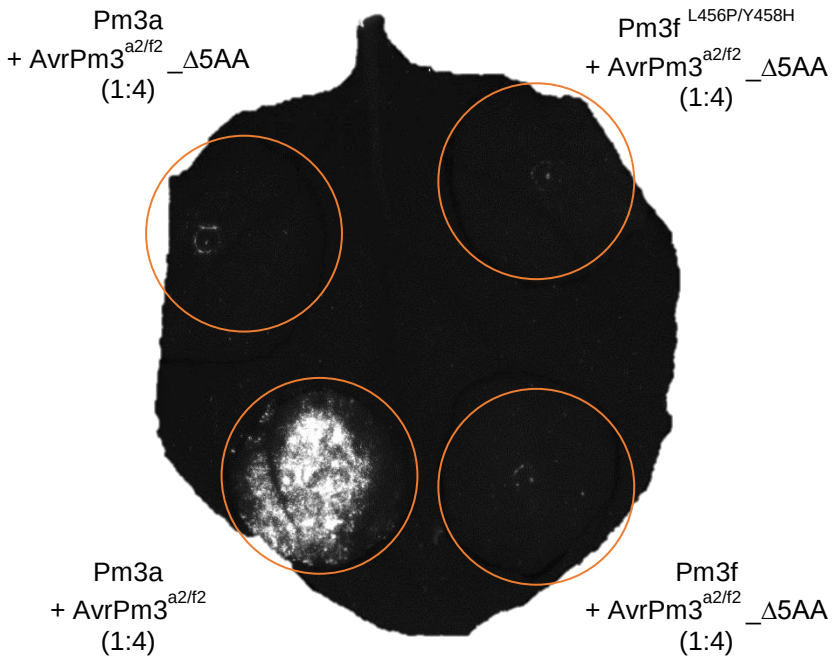


Figure S6. Recognition test of a 5 amino acid deletion of the putative enhancing region (residues 112-116) of AVRPM3$^{A2/F2}$-A. Transient agrobacterium infiltration assays in *Nicotiana benthamiana* show no interaction between the 5 AA deletion mutant (residues 112-116) of AVRPM3$^{A2/F2}$–A and the PM3A or PM3F$^{L456P/Y458H}$ resistance proteins from wheat. Fluorescence imaging of the leaves taken at 5 dpi reveals HR only with the natural AVRPM3$^{A2/F2}$ variant 'A' and the PM3A control.

Supplemental Tables for Chapter 4

Table S1: Primers used for specific amplification of *AvrPm3*$^{a2/f2}$ (pu7) and *SvrPm3*$^{a1/f1}$ (bcg1)

Primer ID	Sequence
pu7_F2	CCTCTGAACCGCCCCATTT
pu7_R2	CAAATCAGGTCACCCCACCA
bcg1_ext_F	ATGGGTTAGCCTTGACATTCC
bcg1_ext_R	AAGACCAGTCATCCCTTCGAC
bcg1_int_F	CATTTGACTTGATTGATGATGATG
bcg1_int_R	CGTTATCCTTGCTTTACCTAGTGA
Y35G_F	AGTGGCAAGTGTCATGACAGGGTCATTGGT
Y35G_R	CTTGCCACTTGAAGCGTTAGCGACAGG
C37G_F	AAGGGTCATGACAGGGTCATTGGTC
C37G_R	ATGACCCTTGTAACTTGAAGCGTTAGCG
H38Q_F	GGACAGGGTCATTGGTCCGGTTAC
H38Q_R	TGACACTTGTAACTTGAAGCGTTAGC
N80S_F	GTGTAATCTTAAAAAGAGGAGAAGAAAATA
N80S_R	TAAAGTATCTTGAACCAAATTTCTGTCCT
N89Y_F	TATATTAAAGTCGAATTTTTTGTGGGAAT
N89Y_R	TTCTTCTCCTCTTTTTAAGATTACATTAA
K91T_F	CAGTCGAATTTTTTGTGGGAAT
K91T_R	TAATATTTTCTTCTCCTCTTTTTAAGATTACATTA
E93K_F	AAATTTTTTGTGGGAATCAACTATT
E93K_R	GACTTTAATATTTTCTTCTCCTCTTTTTAAGATT
K102H_F	TGCACGAAATTATTTATCTACAGGCTTATGTCC
K102H_R	TCGTGCAAATAGTTGATTCCCACAAAAAATTC
K102A_F	TGGCTGAAATTATTTATCTACAGGCTTATGTCC
K102A_R	TCAGCCAAATAGTTGATTCCCACAAAAAATTC

Primer sequences are written in the 5′ to 3′ orientation.

Table S2: Single residue altered AVRPM3$^{A2/F2}$ constructs

Construct	Position	Method
G27S	27	Syn
N31E	31	Syn
S34N	34	Syn
Y35G	35	SDM
C37G	37	SDM
H38Q	38	SDM
H38G	38	Syn
R40L	40	Syn
I42L	42	Syn
T46Y	46	Syn
L47I	47	Syn
N48Q	48	Syn
N48D	48	Syn
D49E	49	Syn
Q50A	50	Syn
I51V	51	Syn
E52D	52	Syn
E52N	52	Syn
K53E	53	Syn
Y55H	55	Syn
Y55F	55	Syn
A58K	58	Syn
E60Q	60	Syn
T63L	63	Syn
P65G	65	Syn
N66R	66	Syn
N66K	66	Syn
L68Y	68	Syn
R69G	69	Syn
R69S	69	Syn
K70P	70	Syn
G71N	71	Syn
Q72E	72	Syn
G75A	75	Syn
S76V	76	Syn
R77G	77	Syn
R77A	77	Syn
Y78T	78	Syn
F79Y	79	Syn
N80S	80	SDM
V81I	81	Syn
I82K	82	Syn
I82Q	82	Syn
L83Y	83	Syn
G86E	86	Bourras et al, 2015
G86D	86	Syn
N89Y	89	SDM
N89I	89	Syn
K91T	91	SDM
K91D	91	Syn
E93K	93	SDM
E93T	93	Syn
E93Q	93	Syn
E93D	93	Syn
F94V	94	Syn
F94I	94	Syn
F95L	95	Haplo
I98F	98	Syn
N99T	99	Syn
Y100I	100	Syn
K102H	102	SDM
K102A	102	SDM
I104V	104	Syn
I105L	105	Syn
Y106W	106	Syn
Y106S	106	Syn
L107V	107	Syn
Q108K	108	Syn
A109V	109	Bourras et al, 2015
V111A	111	Syn
Q112N	112	Syn
S113G	113	Syn
V114Q	114	Syn
L115E	115	Syn
L116I	116	Syn
L116Y	116	Syn
D117E	117	Syn
C118G	118	Syn
Y119E	119	Syn
Y119H	119	Syn
T122R	122	Syn
E123D	123	Syn
R124E	124	Syn
N128S	128	Syn
H132D	132	Syn

Table S3. The worldwide collection of mildew isolates, their AVRPM3$^{A2/F2}$ and SVRPM3$^{A1/F1}$ genotypes, and phenotypes on *Pm3a/f* wheat.

Isolate	Location	Country	AVRPM3$^{A2/F2}$	SVRPM3$^{A1/F1}$	Asosan/ 8*Chancellor (*Pm3a*)	Michigan Amber/ 8*Chancellor (*Pm3f*)	*Formae speciales*	Wheat species collected from
OKS(14)-B-3-2	Oklahoma	USA	A	A	A	V	*tritici*	*Triticum aestivum*
OKH-B-3-3	Oklahoma	USA	A	A	A		*tritici*	*Triticum aestivum*
ARF-C-3-5	Arkansas	USA	A	B	A	V	*tritici*	*Triticum aestivum*
KSM-C-2-5	Kansas	USA	A	B	A	V	*tritici*	*Triticum aestivum*
KSM-C-3-4	Kansas	USA	A	B	A	V	*tritici*	*Triticum aestivum*
KSM-E-1-5	Kansas	USA	A	B	A	V	*tritici*	*Triticum aestivum*
NEL-8	Nebraska	USA	A	B	A	V	*tritici*	*Triticum aestivum*
OKS(14)-A-1-5	Oklahoma	USA	A	B	A	V	*tritici*	*Triticum aestivum*
OKH-A-1-3	Oklahoma	USA	A	B	A		*tritici*	*Triticum aestivum*
OKH-B-1-5	Oklahoma	USA	A	B	A		*tritici*	*Triticum aestivum*
OHW-B-1-2	Ohio	USA	A	C	A	A	*tritici*	*Triticum aestivum*
NEL-5	Nebraska	USA	A		A	V	*tritici*	*Triticum aestivum*
GAD-D-2-4	Georgia	USA	C	A	A		*tritici*	*Triticum aestivum*
FLG-D-2-4	Florida	USA	C	A	I	I	*tritici*	*Triticum aestivum*
NCCL-A-2-5	North Carolina	USA	C	A	I	I	*tritici*	*Triticum aestivum*
GAM-C-2-4	Georgia	USA	C	A	I		*tritici*	*Triticum aestivum*
GAM-D-2-1	Georgia	USA	C	A	I		*tritici*	*Triticum aestivum*
GAP-C-3-2	Georgia	USA	C	A	I		*tritici*	*Triticum aestivum*
MSP1-A-3-2	Mississippi	USA	C	A	I		*tritici*	*Triticum aestivum*
NCC-A-1-3	North Carolina	USA	C	A	I		*tritici*	*Triticum aestivum*
NCC-B-3-1	North Carolina	USA	C	A	I		*tritici*	*Triticum aestivum*
PAC-A-2-5	Pennsylvania	USA	C	A	I		*tritici*	*Triticum aestivum*
PAC-B-3-2	Pennsylvania	USA	C	A	I		*tritici*	*Triticum aestivum*
AK3-11	Arkansas	USA	C	A	V	V	*tritici*	*Triticum aestivum*
FLG-B-2-2	Florida	USA	C	A	V	V	*tritici*	*Triticum aestivum*
KEN4-3	Kentucky	USA	C	A	V	V	*tritici*	*Triticum aestivum*

NYB(14)-B-2-3	New York	USA	C	A	V	V	*tritici*	*Triticum aestivum*
NYG-D-1-1	New York	USA	C	A	V	V	*tritici*	*Triticum aestivum*
NCCL-B-3-1	North Carolina	USA	C	A	V	V	*tritici*	*Triticum aestivum*
NCCL-C-3-5	North Carolina	USA	C	A	V	V	*tritici*	*Triticum aestivum*
PAF(14)-D-1-2	Pennsylvania	USA	C	A	V	V	*tritici*	*Triticum aestivum*
C3-1		USA	C	A	V	V	*tritici*	*Triticum aestivum*
GAM-D-1-4	Georgia	USA	C	A	V		*tritici*	*Triticum aestivum*
MSP1-C-1-2	Mississippi	USA	C	A	V		*tritici*	*Triticum aestivum*
MOB-E-2-5	Missouri	USA	C	A	V		*tritici*	*Triticum aestivum*
MOB-E-3-5	Missouri	USA	C	A	V		*tritici*	*Triticum aestivum*
MOB-D-1-1	Missouri	USA	C	B	V		*tritici*	*Triticum aestivum*
MOB-E-2-2	Missouri	USA	C	B	V		*tritici*	*Triticum aestivum*
FLG-B-3-4	Florida	USA	C		V	V	*tritici*	*Triticum aestivum*
PAF-D-2-2	Pennsylvania	USA	C		V		*tritici*	*Triticum aestivum*
NYB-D-2-2	New York	USA	E	A	I		*tritici*	*Triticum aestivum*
NCC-A-2-1	North Carolina	USA	E	A	I		*tritici*	*Triticum aestivum*
NCF-B-2-1	North Carolina	USA	E	A	I		*tritici*	*Triticum aestivum*
MIR(14)-B-2-1	Mississippi	USA	E	A	V	V	*tritici*	*Triticum aestivum*
NYB(14)-C-2-3	New York	USA	E	A	V	V	*tritici*	*Triticum aestivum*
NYB(14)-D-1-2	New York	USA	E	A	V	V	*tritici*	*Triticum aestivum*
NCM-A-1-1	North Carolina	USA	E	A	V	V	*tritici*	*Triticum aestivum*
OHW-D-3-1	Ohio	USA	E	A	V	V	*tritici*	*Triticum aestivum*
GAD-D-3-1	Georgia	USA	E	A	V		*tritici*	*Triticum aestivum*
NCC-A-3-2	North Carolina	USA	E	A	V		*tritici*	*Triticum aestivum*
NYB-D-3-2	New York	USA	E	A			*tritici*	*Triticum aestivum*
OKS(14)-C-1-3	Oklahoma	USA	E	B	V	V	*tritici*	*Triticum aestivum*
GAP-D-3-3	Georgia	USA	E		I		*tritici*	*Triticum aestivum*
NCF-C-2-5	North Carolina	USA	E		I		*tritici*	*Triticum aestivum*
GAT-B-1-2	Georgia	USA	E		V	V	*tritici*	*Triticum aestivum*
NCM-E-1-2	North Carolina	USA	E		V	V	*tritici*	*Triticum aestivum*
45-10	Gansu province	China	A	A	A	V	*tritici*	*Triticum aestivum*
46-25	Gansu province	China	A	A	A	V	*tritici*	*Triticum aestivum*

46-31	Gansu province	China	A	A	A	V	*tritici*	*Triticum aestivum*
47-3	Gansu province	China	A	A	A	V	*tritici*	*Triticum aestivum*
48-18	Gansu province	China	A	A	A	V	*tritici*	*Triticum aestivum*
SD-5	Shandong province	China	A	A	A	V	*tritici*	*Triticum aestivum*
30-1	Anhui province	China	A	A	A		*tritici*	*Triticum aestivum*
13-50	Guizhou province	China	A	A	A		*tritici*	*Triticum aestivum*
GZ-10	Guizhou province	China	A	A	A		*tritici*	*Triticum aestivum*
GZ-6	Guizhou province	China	A	A	A		*tritici*	*Triticum aestivum*
1-19	Sichuan province	China	A	A	A		*tritici*	*Triticum aestivum*
50-2	Xinjiang province	China	A	A	A		*tritici*	*Triticum aestivum*
9-2	Yunnan province	China	A	A	A		*tritici*	*Triticum aestivum*
36-70	Hebei province	China	A	A	V	V	*tritici*	*Triticum aestivum*
NZ-1	Hubei province	China	A	A	V	V	*tritici*	*Triticum aestivum*
11-99	Hubei province	China	A	A	V		*tritici*	*Triticum aestivum*
9-10	Yunnan province	China	A	A	V		*tritici*	*Triticum aestivum*
6-6	Shannxi province	China	A	D	A	V	*tritici*	*Triticum aestivum*
3-53	Sichuan province	China	A	D	A	V	*tritici*	*Triticum aestivum*
HB-21	Hubei province	China	A	D	A		*tritici*	*Triticum aestivum*
2-25	Sichuan province	China	A	G	A	V	*tritici*	*Triticum aestivum*
24-4	Jiangsu province	China	A	G	A		*tritici*	*Triticum aestivum*
8-9	Yunnan province	China	A	G	A		*tritici*	*Triticum aestivum*
46-30	Gansu province	China	A	G	I	V	*tritici*	*Triticum aestivum*
2-39		China	A		A		*tritici*	*Triticum aestivum*
9-43		China	A		A		*tritici*	*Triticum aestivum*
2-65		China	A		A		*tritici*	*Triticum aestivum*
5-93		China	A		A		*tritici*	*Triticum aestivum*
12-3		China	A		A		*tritici*	*Triticum aestivum*
15-11		China	A		A		*tritici*	*Triticum aestivum*
18-45		China	A		A		*tritici*	*Triticum aestivum*
36-3		China	A		A		*tritici*	*Triticum aestivum*
37-38		China	A		A		*tritici*	*Triticum aestivum*
40-2		China	A		A		*tritici*	*Triticum aestivum*

46-6		China	A		A		*tritici*	*Triticum aestivum*
48-23		China	A		A		*tritici*	*Triticum aestivum*
48-28		China	A		A		*tritici*	*Triticum aestivum*
5-112		China	A		A		*tritici*	*Triticum aestivum*
51-3		China	A		A		*tritici*	*Triticum aestivum*
GZ-12		China	A		A		*tritici*	*Triticum aestivum*
HB-13		China	A		A		*tritici*	*Triticum aestivum*
HB-22		China	A		A		*tritici*	*Triticum aestivum*
NJ-16		China	A		A		*tritici*	*Triticum aestivum*
SC-12		China	A		A		*tritici*	*Triticum aestivum*
SD-3		China	A		A		*tritici*	*Triticum aestivum*
SD-4		China	A		A		*tritici*	*Triticum aestivum*
5-36		China	A		V		*tritici*	*Triticum aestivum*
10-40		China	A		V		*tritici*	*Triticum aestivum*
12-82		China	A		V		*tritici*	*Triticum aestivum*
12-24		China	A		V		*tritici*	*Triticum aestivum*
14-32		China	A		V		*tritici*	*Triticum aestivum*
47-27		China	A		V		*tritici*	*Triticum aestivum*
49-1		China	A		V		*tritici*	*Triticum aestivum*
10-8	Yunnan province	China	B	A	V	V	*tritici*	*Triticum aestivum*
13-51	Guizhou province	China	B	A	V		*tritici*	*Triticum aestivum*
15-17	Henan province	China	B	A	V		*tritici*	*Triticum aestivum*
45-6	Gansu province	China	B	D	V	V	*tritici*	*Triticum aestivum*
6-69	Shannxi province	China	B	D	V	V	*tritici*	*Triticum aestivum*
5-83	Shannxi province	China	B	D	V	V	*tritici*	*Triticum aestivum*
7-8	Shannxi province	China	B	D	V	V	*tritici*	*Triticum aestivum*
1-25	Sichuan province	China	B	D	V	V	*tritici*	*Triticum aestivum*
52-27	Xinjiang province	China	B	D	V	V	*tritici*	*Triticum aestivum*
47-1	Gansu province	China	B	D	V		*tritici*	*Triticum aestivum*
19-11	Jiangsu province	China	B	D	V		*tritici*	*Triticum aestivum*
17-40	Anhui province	China	B	K	A	V	*tritici*	*Triticum aestivum*
14-17	Henan province	China	B	M	A		*tritici*	*Triticum aestivum*

35-18		China	B		A		*tritici*	*Triticum aestivum*
HB-24		China	B		A		*tritici*	*Triticum aestivum*
HG-1		China	B		A		*tritici*	*Triticum aestivum*
1-47		China	B		V		*tritici*	*Triticum aestivum*
11-61		China	B		V		*tritici*	*Triticum aestivum*
1-62		China	B		V		*tritici*	*Triticum aestivum*
3-1		China	B		V		*tritici*	*Triticum aestivum*
18-1		China	B		V		*tritici*	*Triticum aestivum*
21-1		China	B		V		*tritici*	*Triticum aestivum*
21-2		China	B		V		*tritici*	*Triticum aestivum*
2-5		China	B		V		*tritici*	*Triticum aestivum*
21-8		China	B		V		*tritici*	*Triticum aestivum*
28-9		China	B		V		*tritici*	*Triticum aestivum*
18-11		China	B		V		*tritici*	*Triticum aestivum*
6-21		China	B		V		*tritici*	*Triticum aestivum*
11-133		China	B		V		*tritici*	*Triticum aestivum*
13-37		China	B		V		*tritici*	*Triticum aestivum*
13-76		China	B		V		*tritici*	*Triticum aestivum*
18-38		China	B		V		*tritici*	*Triticum aestivum*
35-1		China	B		V		*tritici*	*Triticum aestivum*
35-27		China	B		V		*tritici*	*Triticum aestivum*
37-37		China	B		V		*tritici*	*Triticum aestivum*
39-1		China	B		V		*tritici*	*Triticum aestivum*
39-19		China	B		V		*tritici*	*Triticum aestivum*
39-5		China	B		V		*tritici*	*Triticum aestivum*
E21		China	B		V		*tritici*	*Triticum aestivum*
HB-4		China	B		V		*tritici*	*Triticum aestivum*
E57		China	B				*tritici*	*Triticum aestivum*
12-50	Guizhou province	China	D	A	A	V	*tritici*	*Triticum aestivum*
44-3	Shandong province	China	D	A	V	V	*tritici*	*Triticum aestivum*
34-1	Anhui province	China	D	A	V		*tritici*	*Triticum aestivum*
17-18		China	D		V		*tritici*	*Triticum aestivum*

37-34		China	D		V		*tritici*	*Triticum aestivum*
15-9		China	K		V	V	*tritici*	*Triticum aestivum*
41-5		China	M	D	V	V	*tritici*	*Triticum aestivum*
Isr204	kfar-Masaryk	Israel	A	A	A	A	*tritici*	*Triticum durum*
Isr4	Hula	Israel	A	A	A	V	*tritici*	*Triticum aestivum*
Isr68	Bet Dagan	Israel	A	A	A	V	*tritici*	*Triticum aestivum*
Isr70	Beeri	Israel	A	A	A	V	*tritici*	*Triticum aestivum*
Isr101	Nahal Oz	Israel	A	A	V	V	*tritici*	*Triticum aestivum*
Isr7	Hula	Israel	A	A	V	V	*tritici*	*Triticum aestivum*
Isr94	Ein Hanaziv	Israel	A	A	V	V	*tritici*	*Triticum aestivum*
Isr215	Barkai	Israel	A	A	V	V	*tritici*	*Triticum durum*
Isr92	Tel Aviv	Israel	A		V	V	*tritici*	*Triticum dicoccoides*
Isr13	Hula	Israel	B	A	V	V	*tritici*	*Triticum aestivum*
Isr36	Lahav	Israel	B	A	V	V	*tritici*	*Triticum aestivum*
Isr37	Nahal Oz	Israel	B	A	V	V	*tritici*	*Triticum aestivum*
Isr44	Negev	Israel	B	A	V	V	*tritici*	*Triticum aestivum*
Isr50	Nahal Oz	Israel	B	A	V	V	*tritici*	*Triticum aestivum*
Isr29	Ein Hanaziv	Israel	B	A	V	V	*tritici*	*Triticum aestivum*
Isr16	Nahal Oz	Israel	B	A	V	V	*tritici*	*Triticum durum*
Isr20	Ein Hanaziv	Israel	B	A	V	V	*tritici*	*Triticum durum*
Isr97	Negba	Israel	B	A	V	V	*tritici*	*Triticum durum*
Isr47	Saad	Israel	B		A	V	*tritici*	*Triticum aestivum*
Isr67	Lahav	Israel	B		V	V	*tritici*	*Triticum durum*
Isr205	Kfar-Menahem	Israel	D	A	A	A	*tritici*	*Triticum durum*
Isr214	Akko	Israel	D	A	A	A	*tritici*	*Triticum durum*
Isr208	Gilboa	Israel	D	A	V	A	*tritici*	*Triticum dicoccoides*
Isr106	Nahal Oz	Israel	D	A	V	V	*tritici*	*Triticum aestivum*
Isr107	Nahal Oz	Israel	D	A	V	V	*tritici*	*Triticum aestivum*
Isr216	Ein Shemer	Israel	D	A	V	V	*tritici*	*Triticum aestivum*
Isr217	Kfa Hasidim	Israel	D	A	V	V	*tritici*	*Triticum aestivum*
Isr30P	Talmei Yafe	Israel	D	A	V	V	*tritici*	*Triticum aestivum*
Isr30W	Talmei Yafe	Israel	D	A	V	V	*tritici*	*Triticum aestivum*

Isr43	Yesodot	Israel	D	A	V	V	*tritici*	*Triticum aestivum*
Isr6	Hula	Israel	D	A	V	V	*tritici*	*Triticum aestivum*
Isr96	Negba	Israel	D	A	V	V	*tritici*	*Triticum aestivum*
Isr103	Amiad	Israel	D	A	V	V	*tritici*	*Triticum dicoccoides*
Isr108	Nahal Oz	Israel	D	A	V	V	*tritici*	*Triticum durum*
Isr25	Bet Dagan	Israel	D	A	V	V	*tritici*	*Triticum durum*
Isr8	Hula	Israel	D	A	V	V	*tritici*	*Triticum durum*
Isr52	DirElBalakh	Israel	D	L	V	V	*tritici*	*Triticum durum*
Isr15	Yavor	Israel	D	N	V	V	*tritici*	*Triticum durum*
Isr63	Gilboa	Israel	F	B	A	A	*tritici*	*Triticum dicoccoides*
Isr206	Neve Michael	Israel	F	C	A	A	*tritici*	*Triticum dicoccoides*
Isr201	Tzurit	Israel	F	E	A	A	*tritici*	*Triticum dicoccoides*
Isr207	Bet Meir	Israel	F	E	A	A	*tritici*	*Triticum dicoccoides*
Isr209	K. Revhaya	Israel	F	E	A	A	*tritici*	*Triticum dicoccoides*
Isr210	Givat HaMoreh	Israel	F	E	A	A	*tritici*	*Triticum dicoccoides*
Isr212	Shilat	Israel	F	E	A	A	*tritici*	*Triticum dicoccoides*
Isr220	Bet Oren	Israel	F	E	A	A	*tritici*	*Triticum dicoccoides*
Isr58	Amiad	Israel	F	E	A	A	*tritici*	*Triticum dicoccoides*
Isr66	Amiad	Israel	F	E	A	A	*tritici*	*Triticum dicoccoides*
Isr61	Tabha	Israel	F	E	A	A	*tritici*	*Triticum durum*
Isr203	Misgav	Israel	F	H	A	A	*tritici*	*Triticum dicoccoides*
Isr211	Merav	Israel	F	H	A	A	*tritici*	*Triticum dicoccoides*
Isr109	Tel Aviv	Israel	F	H	A	V	*tritici*	*Triticum dicoccoides*
Isr202	Haifa University	Israel	F		A	A	*tritici*	*Triticum dicoccoides*
Isr9	Hula	Israel	H	A	A	V	*tritici*	*Triticum aestivum*
Isr218	Tel Far	Israel	H	A	V	V	*tritici*	*Triticum aestivum*
Isr33	Erez	Israel	H	A	V	V	*tritici*	*Triticum aestivum*
Isr95	EinHanaziv	Israel	H		V	V	*tritici*	*Triticum durum*
Isr219	Bizaron	Israel	J	A	V	V	*tritici*	*Triticum aestivum*
Isr113	Amiad	Israel	J	A	V	V	*tritici*	*Triticum dicoccoides*
Isr91	Sde Eliahu	Israel	J	A	V	V	*tritici*	*Triticum durum*
Isr1	Hula	Israel	K	A	A	A	*tritici*	*Triticum aestivum*

T1-20	Delley	Switzerland	A	A	A	I	*triticale*	*Triticosecale*
T1-8	Delley	Switzerland	A	A	A	I	*triticale*	*Triticosecale*
T4-16	Guomoëns	Switzerland	A	A	A	I	*triticale*	*Triticosecale*
COPP-2C	Druelle	France	A	A	A	V	*triticale*	*Triticosecale*
T5-12	Ellighausen	Switzerland	A	A	A	V	*triticale*	*Triticosecale*
CAP-39-A1	Cappelle-en-Pévèle	France	A	A	I	V	*triticale*	*Triticosecale*
T4-20	Guomoëns	Switzerland	A	A	I	V	*triticale*	*Triticosecale*
T5-13	Ellighausen	Switzerland	A	A	I	V	*triticale*	*Triticosecale*
T5-9	Ellighausen	Switzerland	A	A	I	V	*triticale*	*Triticosecale*
BAH-2	Bakow	Poland	A	A	V	V	*triticale*	*Triticosecale*
HO-101	Hoogstraten	Belgium	A	A	V	V	*triticale*	*Triticosecale*
THUN-12	Uhnin	Poland	A	A	V	V	*triticale*	*Triticosecale*
T1-23	Delley	Switzerland	A	A	V	V	*triticale*	*Triticosecale*
T6-6	Nyon	Switzerland	A	A	V	V	*triticale*	*Triticosecale*
T5-14	Ellighausen	Switzerland	A	A			*triticale*	*Triticosecale*
T3-4	Reckenholz	Switzerland	A	F	A	A	*triticale*	*Triticosecale*
T3-8	Reckenholz	Switzerland	A	F	A	A	*triticale*	*Triticosecale*
BU-18	Bocholt	Belgium	A	F	A	I	*triticale*	*Triticosecale*
T3-16	Reckenholz	Switzerland	A	F	A	I	*triticale*	*Triticosecale*
T3-9	Reckenholz	Switzerland	A	F	A	V	*triticale*	*Triticosecale*
T4-7	Guomoëns	Switzerland	A	J	A	A	*triticale*	*Triticosecale*
T4-19	Guomoëns	Switzerland	A	J	A	A	*triticale*	*Triticosecale*
07006		Switzerland	A	A	A	V	*tritici*	*Triticum aestivum*
07221		Switzerland	A	A	A	V	*tritici*	*Triticum aestivum*
07237		Switzerland	A	A	A	V	*tritici*	*Triticum aestivum*
07249		Switzerland	A	A	A	V	*tritici*	*Triticum aestivum*
07281		Switzerland	A	A	A	V	*tritici*	*Triticum aestivum*
JIW2		United Kingdom	A	A	A	V	*tritici*	*Triticum aestivum*
JIW48		United Kingdom	A	A	A	V	*tritici*	*Triticum aestivum*
7004		Switzerland	A	A	I	V	*tritici*	*Triticum aestivum*
07296		Switzerland	A	A	V	V	*tritici*	*Triticum aestivum*
07016		Switzerland	A	C	A	A	*tritici*	*Triticum aestivum*
07206		Switzerland	A	C	A	A	*tritici*	*Triticum aestivum*
07214		Switzerland	A	C	A	A	*tritici*	*Triticum aestivum*
07242		Switzerland	A	C	A	A	*tritici*	*Triticum aestivum*
07298		Switzerland	A	C	A	A	*tritici*	*Triticum aestivum*
96224		Switzerland	A	C	A	A	*tritici*	*Triticum aestivum*
97019		Switzerland	A	C	A	A	*tritici*	*Triticum aestivum*

98275	Switzerland	A	C	A	A	*tritici*	*Triticum aestivum*
SYROS95-54	France	A	C	A		*tritici*	*Triticum aestivum*
SYROS2000-15	France	A	C	V	V	*tritici*	*Triticum aestivum*
07286	Switzerland	A	F	A	A	*tritici*	*Triticum aestivum*
96244	Switzerland	A	F	A	A	*tritici*	*Triticum aestivum*
BStone	France	A	F	A	A	*tritici*	*Triticum aestivum*
07017	Switzerland	D	A	V	V	*tritici*	*Triticum aestivum*
WC1110	United Kingdom	G	A	V	V	*tritici*	*Triticum aestivum*
07201	Switzerland	G	B	V	V	*tritici*	*Triticum aestivum*
07203	Switzerland	G	B	V	V	*tritici*	*Triticum aestivum*
07230	Switzerland	G	C	A	V	*tritici*	*Triticum aestivum*
94202	Switzerland	L	A	V	V	*tritici*	*Triticum aestivum*
07010	Switzerland		A	A	V	*tritici*	*Triticum aestivum*
Cun	Australia	A	B	A	V	*tritici*	*Triticum aestivum*
FTA	Australia	A	B	A	V	*tritici*	*Triticum aestivum*
Chikara	Japan	A	A	V	V	*tritici*	*Triticum aestivum*
Shiho	Japan	A	A	V	V	*tritici*	*Triticum aestivum*

Table S4. Phenotypes of progeny from the 96224x94202 population on *Pm1a, Pm4a,* and *Pm4b* wheat.

	Pm3F	Pm1a		Pm4a		Pm4b	
Progeny ID	Michigan Amber 8*Chancellor	Axminster 8*Chancellor	CI_13836	Khapli 8*Chancellor	Khapli 8*Federation	Federation 8*W804	Elysee 8*Federation
201	A	A	nd	A	A	A	A
202	A	A	A	V	V	V	V
205	A	V	V	nd	nd	nd	A
209	A	V	V	A	A	A	A
217	A	A	A	A	A	A	A
221	A	A	A	V	V	V	V
226	A	V	V	A	A	A	A
230	A	A	A	A	A	A	A
241	A	A	nd	V	V	V	V
246	A	A	nd	nd	nd	nd	nd
258	A	A	A	V	V	V	V
260	A	V	V	A	A	A	A
272	A	V	V	A	A	A	A
212	nd	A	A	V	V	V	V
213	nd	A	nd	nd	nd	nd	nd
263	nd	A	A	V	V	V	V
203	V	A	A	A	A	A	A
207	V	A	A	V	V	V	V
210	V	A	nd	V	V	V	V
218	V	V	V	nd	nd	A	A
225	V	V	V	A	A	A	A
231	V	A	A	nd	nd	nd	nd
237	V	A	A	V	V	V	V
240	V	A	A	V	V	V	V
245	V	V	nd	nd	nd	nd	nd
249	V	V	V	A	A	A	A
254	V	A	A	A	A	A	A
257	V	A	A	V	V	V	V
261	V	V	V	A	A	A	A
262	V	A	A	A	A	A	A
264	V	A	A	V	V	V	V
268	V	V	V	A	A	A	A

Virulence (V), avirulence (A), and phenotype not determined (nd) are indicated.

Table S5. Phenotypes of progeny from the 96224x94202 population on *Pm5a* wheat.

	Pm3F	Pm5a
Progeny ID	Michigan Amber/ 8*Chancellor	Hope/ 7*Prins
2	A	A
58	A	V
64	A	V
72	A	V
75	A	V
84	A	A
88	A	V
101	A	V
105	A	V
111	A	I
113	A	V
115	A	V
117	A	A
24	I	A
46	nd	V
130	nd	I
169	nd	V
1	V	A
10	V	V
22	V	V
23	V	A
25	V	V
32	V	V
35	V	V
47	V	A

	Pm3F	Pm5a
Progeny ID	Michigan Amber/ 8*Chancellor	Hope/ 7*Prins
52	V	V
67	V	V
73	V	V
76	V	V
80	V	A
85	V	V
107	V	V
114	V	V
116	V	V
122	V	A
123	V	V
123	V	V
133	V	I
134	V	A
136	V	A
139	v	A
157	V	I
162	V	A
163	v	v
164	V	A
176	V	A
184	V	V
190	V	V
193	V	V
194	V	A

Virulence (V), avirulence (A), intermediate (I) and phenotype not determined (nd) are indicated.

Table S6. Phenotypes of progeny from the 96224x94202 population on *Pm17* wheat.

	Pm3F	Pm17
Progeny ID	Michigan Amber 8*Chancellor	Amigo
56	A	V
63	A	A
72	A	A
75	A	V
83	A	V
84	A	A
87	A	A
158	A	A
159	A	A
161	A	A
168	A	A
170	A	V
179	A	A
183	A	A
187	A	A
198	A	A
199	A	A
71	nd	V
74	nd	V
157	nd	V
61	V	A
65	V	V
70	V	V
76	V	V
80	V	V
82	V	A
85	V	V
86	V	V
139	V	A
140	V	V
141	V	A
143	V	A
152	V	A
162	V	A
163	V	A
165	V	A
166	V	V
167	V	A
169	V	A
171	V	A
172	V	I
173	V	A
174	V	A
175	V	A
176	V	A
177	V	A
182	V	A
184	V	A
185	V	A
186	V	A
188	V	A
189	V	A
192	V	A
193	V	A
197	V	A
200	V	A

Virulence (V), avirulence (A), intermediate (I) and phenotype not determined (nd) are indicated.

Table S7. Student's t-test of significant differences in expression.

		7237 *Avr*	7237 *Svr*	JIW2 *Avr*	JIW2 *Svr*	7004 *Avr*	7004 *Svr*	7296 *Avr*	7296 *Svr*
7237	*Avr*	--	0***	0.1545	0.0008***	0.0104*	0.0275*	0.0028**	0.0001***
7237	*Svr*	--	--	0.0061**	0.0253*	0***	0.0046**	0.0034**	0.0003***
JIW2	*Avr*	--	--	--	0.0051**	0.0818	0.0481*	0.015*	0.0186*
JIW2	*Svr*	--	--	--	--	0.0017*	0.0036**	0.0053**	0.0071**
7004	*Avr*	--	--	--	--	--	0.021*	0.0073**	0.0003***
7004	*Svr*	--	--	--	--	--	--	0.0071**	0.0093**
7296	*Avr*	--	--	--	--	--	--	--	0.4186
7296	*Svr*	--	--	--	--	--	--	--	--

*, **, and *** indicate p-values <0.05, 0.01, and 0.001, respectively

Table S8. Nucleotide polymorphisms in the *AvrPm3*$^{a2/f2}$ haplotypes.

AvrPm3a2/f2 Haplotype	62	71	72	73	74	75	76	77	80	87	91	93	114	154	156	175	198	207	239	257	265	272	277	279	285	326	355	365	369	%id
***A*[a]**	**C**	**G**	**-**	**-**	**-**	**-**	**-**	**-**	**G**	**C**	**A**	**C**	**T**	**G**	**A**	**C**	**C**	**G**	**A**	**G**	**A**	**A**	**G**	**A**	**T**	**C**	**T**	**C**	**G**	
A-T	.	.	-	-	-	-	-	-	.	.	.	.	.	.	.	T	.	.	.	.	.	.	.	.	.	.	.	.	.	0.9975
C	.	.	-	-	-	-	-	-	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	G	.	.	.	.	0.9975
D	.	.	-	-	-	-	-	-	.	.	.	.	.	.	.	.	A	.	G	.	.	C	A	.	.	.	.	.	.	0.9898
F	.	.	-	-	-	-	-	-	.	.	.	.	.	.	.	.	A	T	.	A	.	.	.	.	.	.	.	.	.	0.9924
G	T	A	C	T	C	G	C	C	T	T	G	A	G	.	.	.	A	.	.	A	T	.	.	.	.	T	C	.	.	0.9549
M	.	.	-	-	-	-	-	-	.	.	G	A	G	A	T	.	.	.	.	A	.	.	.	.	.	.	.	.	C	0.9822
B	.	.	-	-	-	-	-	-	.	.	.	.	G	.	.	.	.	.	.	A	.	.	.	.	.	.	.	.	C	0.9924
K	.	.	-	-	-	-	-	-	.	.	.	.	G	.	.	.	.	.	.	A	.	.	.	.	.	.	.	.	.	0.9949
H	.	.	-	-	-	-	-	-	.	.	.	.	.	.	.	.	.	.	.	A	.	.	.	.	.	.	.	.	C	0.9949
J	.	.	-	-	-	-	-	-	.	.	.	.	.	.	.	.	.	.	.	A	.	.	.	.	.	.	.	G	C	0.9924
L[a]	.	.	-	-	-	-	-	-	.	.	.	.	.	.	.	.	.	.	.	A	.	.	.	.	.	T	.	.	.	0.9949
E	.	.	-	-	-	-	-	-	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	T	.	.	.	.	.	0.9975

Nucleotide polymorphisms compared to the avirulent haplotype are depicted. Dashes indicate gaps and dots indicate identical residues. The 'A-T' haplotype was found only in *B.g. triticale* isolates and contains the only synonymous mutation in the haplotypes.

[a] Bourras et al, 2015

Note S1. Genetic evidence indicating *AvrPm3$^{a2/f2}$* is not recognized by *Pm1a, Pm2, Pm3b-e, Pm3g, Pm4a, Pm4b, Pm5a, Pm8,* and *Pm17.*

In order to assess if *AvrPm3$^{a2/f2}$* is genetically controlling avirulence to other *R* genes, we compared phenotype segregation data on *Pm3f* to those on *Pm1a*, *Pm4a*, *Pm4b*, and *Pm5a* in a subset of haploid progeny originating from a cross between the mildew isolates 96224 and JIW2 (Parlange et al., 2015, Supporting Information Table S4 and S5). If avirulence on *Pm1a*, *Pm4a*, *Pm4b*, or *Pm5a* is controlled by the same genetic locus as *Pm3f*, we expect that the progeny will have the same segregation pattern on *Pm3f* as compared to *Pm1a*, *Pm4a*, *Pm4b*, or *Pm5a*. For *Pm1a*, *Pm4a*, and *Pm4b*, 32 progeny were tested and the phenotype segregation pattern were completely different to those on *Pm3f* (Supporting Information Table S4). For the phenotypes on *Pm5a*, 50 progeny were tested, and phenotype segregation was also completely different to that on *Pm3f* (Supporting Information Table S5). This data indicates that avirulence on *Pm3f* is genetically unlinked to avirulence on *Pm1a, Pm4a, Pm4b, and Pm5a*. Therefore, it is unlikely that *AvrPm3$^{a2/f2}$* is also recognized by *Pm1a, Pm4a, Pm4b,* or *Pm5a*.

For *Pm17*, we used the same approach in a subset of 56 haploid progeny originating from a cross between the mildew isolates 96224 and 94202 (Bourras et al. 2015; Supporting Information Table S6). Here also, phenotype segregation on *Pm17* and *Pm3f* were completely different, indicating that avirulence on *Pm3f* is genetically unlinked to avirulence on *Pm17*. Therefore, it is unlikely that *AvrPm3$^{a2/f2}$* is also recognized by *Pm17*.

To conclude, phenotype segregation data from Parlange et al. (2015), Bourras et al. (2015), and Supporting Information Tables S4-S6, as well as genetic mapping data from Parlange et al. (2015), Bourras et al. (2015), and Praz et al. (2017), indicate that *AvrPm3$^{a2/f2}$* is not controlling avirulence towards *Pm1a*, *Pm2*, *Pm3b-e*, *Pm3g*, *Pm4a*, *Pm4b*, *Pm5a*, *Pm8*, and *Pm17*.

Appendix III - Supplemental Table for Chapter 5

Supplemental table 1: Tajima's D test of neutrality (*D*) for both *AvrPm3*$^{a2/f2}$ and *SvrPm3*$^{a1/f1}$.

Gene	Origin	m	S	p_s	Θ	π	D
AvrPm3$^{a2/f2}$	USA	56	5	0.011	0.002	0.422	1.678
AvrPm3$^{a2/f2}$	Europe	46	7	0.018	0.004	0.001	-2.039
AvrPm3$^{a2/f2}$	Israel	46	8	0.02	0.005	0.009	2.482
AvrPm3$^{a2/f2}$	China	101	11	0.028	0.005	0.005	-0.224
AvrPm3$^{a2/f2}$	Global	253	16	0.041	0.007	0.006	-0.428
SvrPm3$^{a1/f1}$	USA	49	3	0.007	0.002	0.002	0.38
SvrPm3$^{a1/f1}$	Europe	51	6	0.014	0.003	0.004	1.13
SvrPm3$^{a1/f1}$	Israel	47	6	0.014	0.003	0.002	-0.953
SvrPm3$^{a1/f1}$	China	41	4	0.009	0.002	0.004	1.651
SvrPm3$^{a1/f1}$	Global	192	11	0.025	0.004	0.003	-0.569

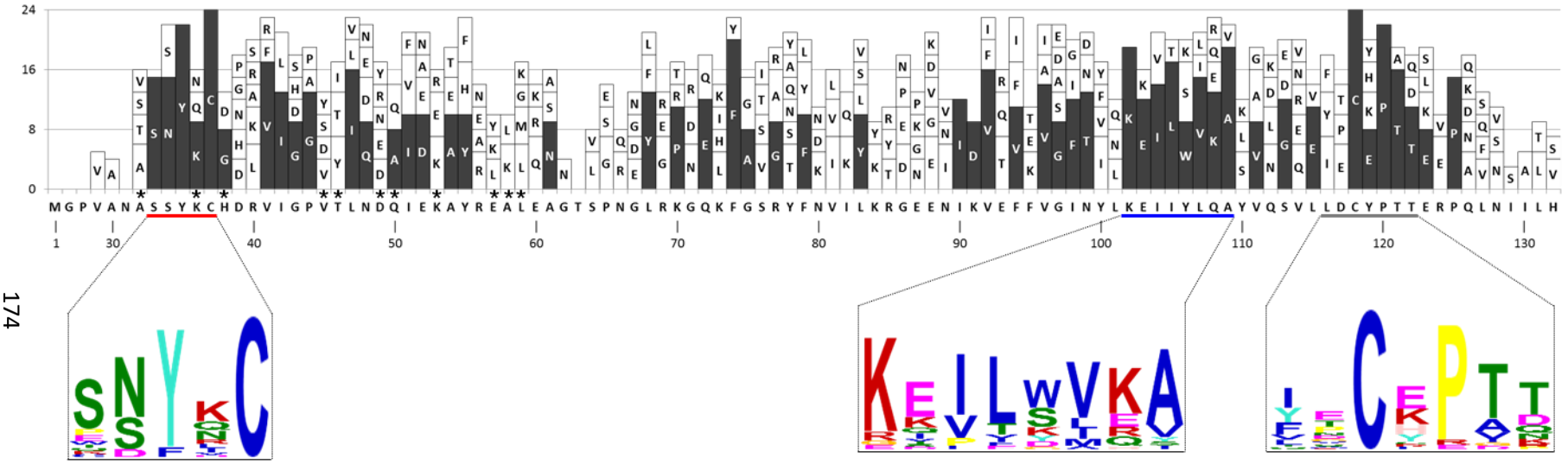
Codon positions included were 1st+2nd+3rd+Noncoding. All positions containing gaps and missing data were eliminated. Evolutionary analyses were conducted in MEGA6 [2]. Abbreviations: *m* = number of sequences, *S* = number of segregating sites, p_s = S/n, π = nucleotide diversity, and *D* is the Tajima test statistic (Tajima, 1989; Tamura et al, 2013; Nei and Kumar, 2000).

Appendix IV - *AvrPm3$^{a2/f2}$* structural modeling studies

Modelling the *AvrPm3$^{a2/f2}$* effector family reveals a putative intein-like structure

The AVRPM3$^{A2/F2}$ family shares no homology to any known proteins, but internally they share conservation of residues that are normally implicated in the formation of secondary structure, such as hydrophobic residues or cysteines (Appendix Figure A.1). Therefore, we postulated that a common tertiary structure prediction for all family members would be a good indication of a likely structural model for AVRPM3$^{A2/F2}$. We submitted the mature protein sequences of all family members without signal peptides to the RaptorX Structure Prediction platform (Källberg et al, 2012), which uses a protein threading algorithm to predict homology based on protein fold and not sequence homology. Despite overall low significance levels, the only template predicted to share homology with more than one family member was shared by more than half the family (13/25) and constituted 13 of the top 16 most significant predictions (Appendix Table A.1). This common predicted template structure is an intein homing endonuclease from *Thermococcus kodakarensis* (PDB: 2cw7A), and the family aligns specifically to the self-excising intein domain (Appendix Figure A.2A; Matsumura et al, 2006). This structure is congruent with the secondary structure prediction from Jpred4 based on an alignment of the family (Appendix Figure A.3). Inteins are internal protein elements capable of self-splicing post-translationally to form the mature protein product, and they are found in all three domains of life (Novikova et al, 2014). The intein homing endonuclease structure was not a predicted template for AVRPM3$^{A2/F2}$ modelling by RaptorX. Therefore we used the predicted intein structure of the closest relative, *Pu_24* (58% aa identity), as a template onto which we modelled the AVRPM3$^{A2/F2}$ protein sequence using Phyre2 (Kelley et al, 2015; Appendix Figure A.2B).

Using this model we then compared the results of the mutagenesis study and the proposed intein structure of AVRPM3$^{A2/F2}$ (Appendix Figure A.4A-C). The residues under diversifying selection in the family, which might therefore represent residues important for effector function but not recognition, correspond to a pocket formed by a loop and the α-helix at the N-terminal region of the protein (residues 32-59; Appendix Figure A.4A-C, green residues). This region is overlapping with the YxC motif (orange residues). In the mutagenesis screen, recognition was particularly insensitive to mutations in the region from residue 60-76, which in this model corresponds to the proximal loop following the single α-helix. The region from residue 86-95 contains the highest concentration of haplotype mutations (5 over 9 sites) and corresponds to the next loop and the subsequent β-strand that is highly exposed in the protein model. All *AvrPm3$^{a2/f2}$* haplotypes other than the recognized *AvrPm3$^{a2/f2}$-A* encode at least one disrupting mutation in this region.



Appendix Figure A.1. Conserved residues and motifs within the *AvrPm3*$^{a2/f2}$ family. Protein sequences of 24 family members without the signal peptide were aligned and compared to the *AvrPm3*$^{a2/f2}$ sequence. Residues present in at least three family members (white histograms) and residues conserved in more than 8 of 24 family members (black histograms) are shown. Asterisks indicate residues under diversifying selection (probability of (ω>1) >0.95). The SNYxC motif (red line), as well as the novel KxILxVKA (blue line) and I/F/YxCxPTT motif (green line) are depicted.

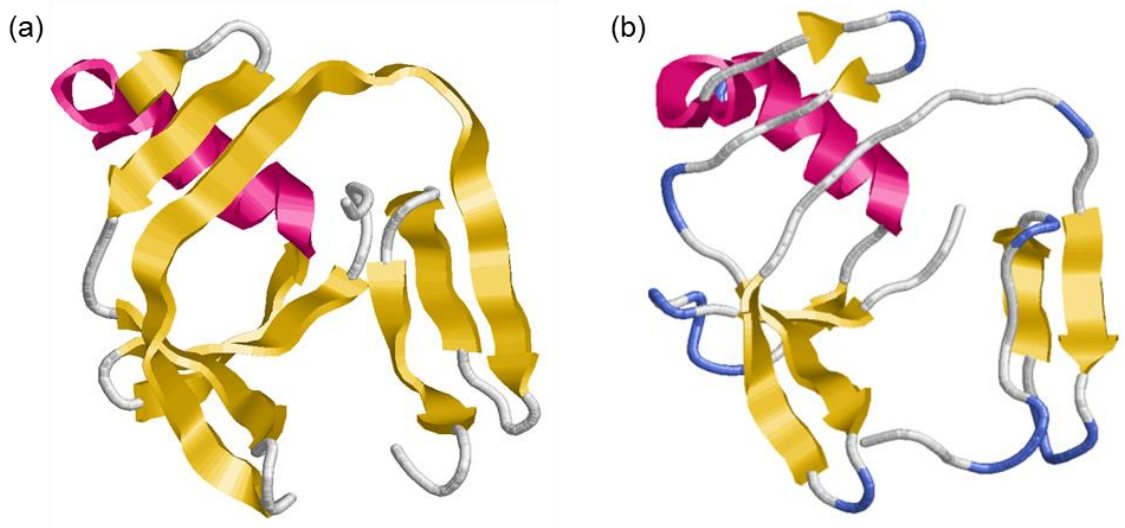
Appendix Table A.1. RaptorX structure predictions for the *AvrPm3*$^{a2/f2}$ family.

pu7 Family	Best template	p-value	uGDT (GDT)	Expression
Bgt_avrF2_pu28_p	2cw7A	1.50E-03	41 (41)	high
Bgt_avrF2_pu-12_p	2cw7A	1.54E-03	39 (37)	
Bgt_avrF2_pu-9_p	2cw7A	1.90E-03	38 (38)	high
Bgt_avrF2_pu-5_p	2cw7A	2.19E-03	38 (36)	
Bgt_avrF2_pu12B_p	2cw7A	2.39E-03	38 (37)	
Bgt_avrF2_pu-14_p	2cw7A	2.59E-03	39 (37)	
Bgt_avrF2_pu-3_p	2cw7A	2.71E-03	38 (38)	low
Bgt_avrF2_pu-24_p	2cw7A	3.16E-03	38 (35)	
Bgt_avrF2_pu-13_p	2cw7A	3.24E-03	38 (37)	
Bgt_avrF2_pu-21_p	2hv2A	4.14E-03	34 (31)	
Bgt_avrF2_pu-6_p	2cw7A	4.19E-03	41 (38)	low
Bgt_avrF2_pu-20_p	2cw7A	4.64E-03	39 (35)	
Bgt_avrF2_pu-8_p	2wzpP	4.93E-03	30 (32)	low
Bgt_avrF2_pu-19_p	2cw7A	5.93E-03	34 (33)	
Bgt_avrF2_pu-1_p	4khaA	6.86E-03	35 (36)	high
Bgt_avrF2_pu-2_p	2cw7A	7.05E-03	36 (36)	
Bgt_avrF2_pu-23_p	4pa0A	7.10E-03	39 (35)	low
Bgt_avrF2_pu-15_p	5btuA	7.77E-03	33 (35)	low
Bgt_avrF2_pu-25_p	3pgbA	7.89E-03	31 (32)	
AvrPm3a2/f2	3qtdA	8.38E-03	33 (31)	high
Bgt_avrF2_pu-10_p	1yqfA	9.80E-03	39 (36)	
Bgt_avrF2_pu-4_p	3cygA	1.17E-02	30 (28)	
Bgt_avrF2_pu-27_p	3ozqA	1.18E-02	26 (27)	high
Bgt_avrF2_pu-11_p	3qsqA	1.29E-02	23 (26)	
Bgt-20382_pu-26_p	1s96A	1.96E-02	27 (28)	low

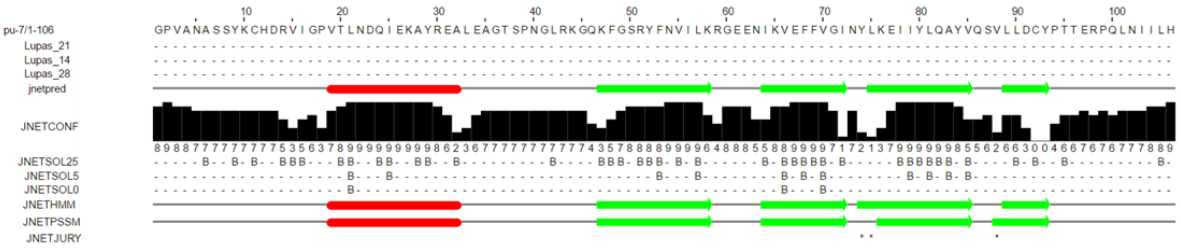
Protein sequences of the *AvrPm3*$^{a2/f2}$ family without signal peptide were individually submitted to the RaptorX Structure Prediction platform and the results are summarized here. The results are listed by significance of the predicted homology. The 'Best template' column lists the name of the PDB file used as the main template for the predicted structure. The 2cw7A structure is a homing endonuclease containing an intein domain. It is the most commonly chosen template for the *AvrPm3*$^{a2/f2}$ family. The uGDT(GDT) is a measure of strength of the proposed model. The five highest and five lowest expressed family members are indicated.

The second and third motifs from the family (KxILxVKA, red; and I/Y/FxCxPTT, blue) are partially included in, and partially flank, the sites most affecting recognition specificity according to the mutagenesis results (residues 112-116, teal). Taken together, this region of two β-strands and a connecting loop (residues 102-122) constitutes the final highly exposed loop region of the intein-based structural model.

Although the significance of the predicted structural homology between individual family members and the intein domain is low, several congruent evidences support a hypothesis of an intein-like model for the *AvrPm3*$^{a2/f2}$ family. First, the prediction was shared by a majority of the members of the family, which share limited homology outside conserved structural residues. The model also conforms to the secondary structure prediction of an alignment of the family from Jpred4 that calculated the significant likelihood of an N-terminal α-helix followed by a gap and then 3-4 subsequent β-sheets (Appendix Figure A.3). Second, the motifs and residues under positive selection in the family correspond to structural domains in the predicted protein model such as a pocket formed by the exposed residues of the α-helix. And finally, the distinct domains identified by the mutagenesis screen overlap with structural domains such as loops formed by antiparallel β-sheets separated by a short turn (Appendix Figure A.4), constituting a highly exposed region that might be important for direct interaction with *Pm3a/f*. Therefore, the proposed intein structure provides a solid basis onto which we can base hypotheses and test the role of individual residues in conferring recognition specificity.



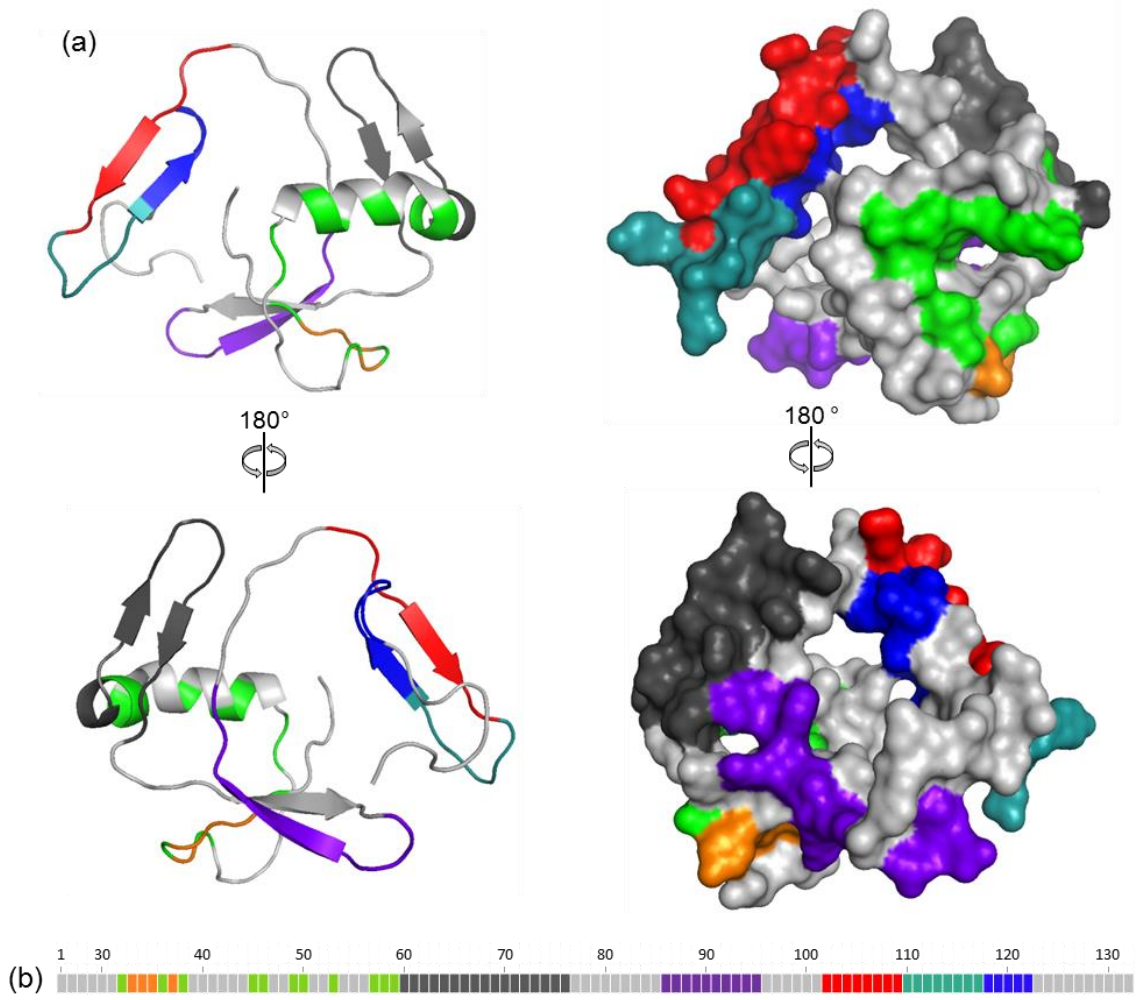
Appendix Figure A.2. Structural model of *AvrPm3*$^{a2/f2}$ based on the intein domain of a homing endonuclease. (a) The crystal structure of the partial intein domain of the homing endonuclease from *Thermococcus kodakarensis* (residues 1-105; PDB: 2cw7A) that shares structural homology with 13 *AvrPm3*$^{a2/f2}$ family members. Residues are colored according to structure (α-helix, pink; β-strands, yellow). (b) The protein sequence of *AvrPm3*$^{a2/f2}$ was modeled onto the predicted intein structure of *Pu_24*, the closest relative predicted to share structural homology to an intein domain. Residues are colored according to structure (α-helix, pink; turns, blue; β-strands, yellow).



Appendix Figure A.3. Jpred4 secondary structure prediction based on an alignment of the *AvrPm3*$^{a2/f2}$ family. An alignment of the *AvrPm3*$^{a2/f2}$ effector family without the signal peptide sequences was submitted to the Jpred4 secondary structure prediction server (Jnet version: 2.3.1). The sequence of *AvrPm3*$^{a2/f2}$ (pu-7, 1-106 residues) is shown at the top. Lupas_21, Lupas_14, and Lupas_28 are binary predictions of coiled-coil structure. Jnetpred is the consensus prediction, with helices marked in red and sheets in green. JNetCONF is the confidence estimate for the prediction: high values indicate high confidence. JNETSOL25, JNETSOL5, and JNETSOL0 are binary solvent accessibility predictions at 24%, 5% or 0% solvent accessibility (B=burried). JNETHMM is predicts helices (red) and sheets (green) based on Hidden Markov Model profiles. JNETPSSM uses a position-specific scoring matrix to predict helices (red) and sheets (green). JNETJURY indicates (*) where JNETJURY was invoked to rationalize significantly different primary predictions from other annotations.

Mutations toward enhancing or eliminating putative intein enzymatic function abolish recognition by *Pm3a/f*

Inteins are characterized by short, specific motifs necessary for enzymatic autoproteolysis. One such motif present in all inteins is TxxH, which is found near the middle of the protein and is crucial for intermediate biochemical steps during autoproteolysis (Shah and Muir, 2014). In the effector family we find in a similar location, a highly conserved TxxK (13/24 family members) at residues 99-103. At this location *AvrPm3*$^{a2/f2}$ encodes NxxK. It is common for an Asp to replace the Thr in this motif, as they have similar biochemical properties. Though rarer, the use of Lys in place of a His residue in the TxxH motif has been previously described (Tori et al, 2012). To test whether a putative enzymatic reaction at this motif affects recognition specificity by PM3A/F, we mutated these residues independently and in combination toward the common TxxH motif (N99T/K103H), and also replaced the catalytic Lys by Ala (K103A). All constructs were coinfiltrated with PM3A, PM3F, and PM3F$^{L456P/Y458H}$ in *N. benthamiana* and HR was assessed after 5 days. All constructs lost the cell death response with all PM3 constructs. These results suggest that putatively enhancing or eliminating the potential enzymatic activity in these residues results in disruption of recognition by PM3A/F. Also, it is unlikely that any putative enzymatic activity is responsible for the recognition response, as both enhancing and eliminating it have the same outcome on recognition.



Appendix Figure A.4. Overlay of functionally important residues in *AvrPm3*$^{a2/f2}$ onto the predicted intein structural model. (A) The predicted intein structure of *AvrPm3*$^{a2/f2}$ in cartoon (left) and space-filled (right) with (B) an overview of the domains identified in the mutagenesis screen. Residues predicted to be under positive selection in the family (green) and the region containing the most natural diversity (purple) are shown. The three motifs common among the family are indicated where not overlapping with other domains (SNYxC, orange; KxILxVKA, red; and I/F/YxCxPTT, blue). The region containing most of the residues affecting the strength of the HR and recognition specificity is colored teal.